(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 January 2001 (25.01.2001)

PCT

(10) International Publication Number
WO 01/05825 A2

(51) International Patent Classification[7]: C07K 14/00

(21) International Application Number: PCT/US00/19429

(22) International Filing Date: 17 July 2000 (17.07.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/354,881 16 July 1999 (16.07.1999) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 09/354,881 (CIP)
Filed on Not furnished

(71) Applicant *(for all designated States except US)*: **HYSEQ, INC.** [US/US]; 670 Almanor Avenue, Sunnyvale, CA 94086 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BALLINGER, Dennis, G.** [US/US]; 38 Bishop Lane, Menlo Park, CA 90424 (US). **MONTGOMERY, Julie, Reeder** [US/US]; 23411 Glenwood Drive, Los Gatos, CA 95033 (US).

(74) Agent: **WILLIAMS, Joseph, A., Jr.**; Marshall, O'Toole, Gerstein, Murray & Borun, 6300 Sears Tower, 233 S. Wacker Drive, Chicago, IL 60606-6402 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Fd prior to Nov 29, 2000 :: not 102(e)

WO 01/05825 A2

(54) Title: NOVEL ANGIOPOIETIN MATERIALS AND METHODS

(57) Abstract: The present invention provides novel nucleic acids encoding human angiopoietins, the novel polypeptides encoded by these nucleic acids and uses of these and related products.

NOVEL ANGIOPOIETIN MATERIALS AND METHODS

This application is a continuation-in-part of U.S. patent application Serial No: 09/354,881, filed July 16, 1999.

FIELD OF THE INVENTION

The present invention related to novel polynucleotides encoding human angiopoietin polypeptides, along with therapeutic, diagnostic and research utilities thereof.

BACKGROUND

Vascular development and neovascularization are regulated in part by paracrine signals transduced by transmembrane tyrosine receptor kinases (TRKs), or kinase receptors, on endothelial cells. One family of TRKs necessary for vascular development in the embryo and neovascularization in adults comprises the receptors flk-1, flt-4, and flt-1 which interact with the cytokine vascular endothelial growth factor (VEGF) [Mustonen and Alitalo, J. Cell. Biol. 129:895-898 (1995); Joukov, et al., EMBO J. 15:290-298 (1996)]. A second family of TRKs comprises endothelial-specific transmembrane tyrosine kinases with immunoglobulin and epidermal growth factor domains, including Tie-1 and Tie-2 [Dumont, et al., Dev. Dyn. 203:80-92 (1995); Maisonpierre, et al., Oncogene 8:1631-1637 (1993); Sato, et al., Proc. Natl. Acad. Sci. (USA) 90:9355-9358 (1993); Ziegler, et al., Oncogene 8:663-670 (1997); Dumont, et al., Oncogene 7:1471-1480 (1992); Schnurch and Rizzu, Development 119:957-968 (1993) Sato, et al., Nature 378:70-74 (1995)]. To date, no physiological ligand has been identified that interacts with Tie-1, while the natural ligands for Tie-2 comprise a family of proteins including the angiopoietins. Two distinct angiopoietin proteins, Ang-1 [Davis, et al., cell 87:1161-1169 (1996)] and Ang-2 [Maisonpierre, et al., Science 277:55-80 (1997)] have been identified that interact with Tie-2 *in vivo*. Two additional angiopoietins, Ang-3 and Ang-4 [Valenzuela, et al., Proc. Natl. Acad. Sci. (USA) 96:1904-1909 (1999)], have been isolated and, at least when expressed as recombinant chimeric proteins, bind Tie-2 *in vitro*.

Expression of Tie-2 is critical to formation of embryonic vasculature as demonstrated in mice deficient in Tie-2 expression that display a lethal phenotype with generalized defects in vascular structure [Dumont et al., Genes Dev. 8:1897-1909 (1994)]. In Tie-2-deficient mice, abnormally rounded endothelial cells are detected, indicating a failure of endothelial/matrix interaction [Patan, Microvascular Res. 56:1-21 (1998)]. Rounded endothelial cells do not spread or flatten normally and do not associate with periendothelial cells [Witzenbichler, et al., J. Biol. Chem. 273:13514-13521 (1998)]. Reduced interaction of the endothelial with the extracellular matrix causes collapse of the sinus venous and other larger and smaller vessels, and occlusion of the connection between atrium and ventricle, as well as ventricle and aorta [Patan, supra]. This failure, along with a failure to recruit mesenchymal cells to form periendothelial cells, indicates the importance of the cell/matrix and cell/cell interactions for maintaining the configuration of the normal lumen [Patan, supra]. In normal adult quiescent vessels, the periendothelial cells (pericytes) constitutively secrete Ang-1, which enhances contact between neighboring endothelial cells and between periendothelial cells [Tsurumi, et al., J. Clin. Invest. 93:662-670 (1994)] thereby maintaining endothelial integrity and orientation of endothelial cells on basal lamina [Witzenbichler,, et al., supra]. Disruption of the endothelial monolayer may lead to upregulation of Ang-1 expression in surrounding cells, or lead to changes in Tie-2 receptor expression and subsequent reendothelialization of the denuded areas [Witzenbichler,, et al., supra]. It has been proposed that expression of Ang-1 supplements VEGF in recruiting endothelial cells and promoting reendothelialization. [Witzenbichler, et al., supra]. In addition, it has been suggested that, under conditions of postnatal angiogenesis, such as tissue ischemia and tumor growth, the action of Ang-1 on endothelial cells may be important for initiation of new capillary sprouting, as well as the movement of endothelial cells toward each other, an activity that is required for fusion into capillary structures [Witzenbichler, et al., supra].

In general, angiopoietin polypeptides comprise three predominant domains [Davis, et al., Cell 87:1161-1169 (1996)]. At the amino terminus is a distinctive region in each protein that shows no homology to other known proteins. Adjacent this region is an alpha helix-rich domain that is common to proteins that tend to multimerize. In fact, it is

believed that active angiopoietins act as multimeric aggregates, perhaps as heteromultimers, comprising several different angiopoietins. The distinctive amino terminal domain and the alpha helix regions from different angiopoietins can be substituted between proteins, but the signal transduction capacity of the chimeric protein is dictated by the third protein region, designated the fibrinogen-like domain (FD). The FD region of the angiopoietins comprises receptor binding sequences and dictates whether the protein is an agonist or an antagonist of Tie-2 signal transduction. For example, despite 60% amino acid sequence homology between Ang-1 and Ang-2, Ang-1 is an agonist of Tie-2 signaling on endothelial cells, which upon binding induces Tie-2 autophosphorylation, while Ang2 does not transduce Tie-2 signalling and competitively inhibits receptor autophosphorylation. Interestingly, however, in the rare instances in which Tie-2 is expressed on cells other than endothelial cells, both Ang-1 and Ang-2 are agonists for Tie-2 signaling. It is therefore believed that the ratio of Ang-1:Ang-2 regulates vessel maturation and stabilization, and that elevated levels of Ang-2 lead to blood vessel destabilization and subsequent regression of the vasculature, as demonstrated during follicle atresia and corpus luteum regression (luteolysis) in the cyclic ovary [Goede et al., Lab. Invest. 78:1385-1394 (1998)].

Ang-1 has also been shown to prevent cell death in HUVEC cells in vitro [Papapetropoulos, et al., Lab. Invest. 79:213-233 (1999)], as well as to promote in vitro differentiation of aorta-gonad-mesonephros cells into hemangioblasts, the progenitors of both hematopoietic and endothelial cells [Hamaguchi, et al., Blood 93:1549-1556 (1999)]. Overexpression of Ang-1 in the skin of transgenic mice increases the extent of vascularization [Suri, et al., Science 282:468-471 (1998)]. Overexpression of Ang-2 in transgenic mice, however, mimics the phenotype of the loss of Ang-1 expression [Maisonpierre, et al., Science 277:48-50 (1997)], confirming the antagonistic effects of these Tie-2 ligands. There is evidence that local expression of Ang-2, in conjunction with vascular-endothelial growth factor, can promote angiogenesis [Stratmann, et al., Am. J. Pathol. 153: 1459-1466 (1998)].

Previous results [Koblizek, et al., Curr. Biol. 8:529-32 (1998)] using a monolayer of endothelial cells, cultured on microcarrier beads and embedded in three-dimensional fibrin gels, showed that recombinant Ang-1 induced the formation of

capillary sprouts in a manner that was completely inhibited by soluble Tie-2. In contrast to VEGF, Ang-1 was only very weakly mitogenic for endothelial cells. However, VEGF and Ang-1 acted synergistically to induce sprout formation. The data suggest that vessel formation requires a cascade of activity in which VEGF and angiopoietins, along with their receptors, are important regulators.

Thus, there is a great need for identification of angiopoietins which may be useful for modulating vascular stability and neovascularization associated with various pathologies. Identification of angiopoietin species permits the identification of compounds that can modulate biological activity of specific members of the angiopoietin family, and/or more than one member of the angiopoietin family wherein the members share one or more biological activities. Knowledge of angiopoietins, the genes encoding them, and modulators of their biological activity permit development of therapeutic treatments for conditions, and in particular pathologies, associated with aberrant angiopoietin activity, and as well as methods to augment angiopoietin activity which may increase or decrease angiogenesis.

SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, in particular, novel human angiopoietin proteins and active variants thereof, isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide

encoding a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48. The isolated polynucleotides of the invention further include, but are not limited to, a polynucleotide comprising the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47; a polynucleotide comprising the full length protein coding sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47; and a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47. The polynucleotides of the present invention also include, but are not limited to, polynucleotides that encode polypeptides with angiopoietin activity and that hybridize under stringent hybridization conditions to the complement of (a) the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47, or (b) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48; a polynucleotide which is an allelic variant of any polynucleotide recited above; a polynucleotide which encodes a species homolog of any of the proteins recited above; or a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptide having an amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48. The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48 or a portion thereof corresponding to the full length or mature protein. Polypeptides of the invention also include polypeptides with angiopoietin activity that are encoded by (a) polynucleotides set out in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of the angiopoietin protein sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48 and "substantial equivalents" thereof (e.g., with 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that retain angiopoietin activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also relates to methods for producing polypeptides of the invention comprising growing a culture of the cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the protein from the cells or the culture medium in which the cells are grown. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect or quantify the presence of the particular cell or tissue mRNA in a sample using, e.g., in situ hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

In particular, where the polypeptide has anti-angiogenic activity, for example, anti-Ang-1-like activity, or Tie-2 antagonist activity, the polypeptides and polynucleotides of the invention can be utilized, for example, as part of methods for the prevention and/or treatment of angiopoietin mediated disorders including disorders involving hypervascularization, often associated with tumorogenesis, or any of the disorders described below. Where the polypeptide promotes angiogenesis, for example, Ang-1-like activity, or Tie-2 agonist activity, polypeptides and polynucleotides can be utilized, for example, as part of treatment for disorders that would benefit from increased vascularization, for example wound healing, osteonecrosis, and any of the other disorders described herein.

The methods of the present invention further relate to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited herein.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compounds that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention.

The methods of the invention also include methods for the treatment of disorders as recited above which may involve the administration of such compounds to individuals exhibiting symptoms or tendencies related to disorders as recited herein. In addition, the invention encompasses methods for treating diseases or disorders as recited herein comprising the step of administering compounds and other substances that modulate the

overall activity of the target gene products. Compounds and other substances can effect such modulation either on the level of target gene expression or target protein activity.

DETAILED DESCRIPTION OF THE INVENTION

1. DEFINITIONS

The term "nucleotide sequence" refers to a heteropolymer of nucleotides or the sequence of these nucleotides. The terms "nucleic acid" and "polynucleotide" are also used interchangeably herein to refer to a heteropolymer of nucleotides. Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" is a stretch of polypeptide nucleotide residues which is long enough to use in polymerase chain reaction (PCR) or various hybridization procedures to identify or amplify identical or related parts of mRNA or DNA molecules.

The terms "oligonucleotides" or "nucleic acid probes" are prepared based on the polynucleotide sequences provided in the present invention. Oligonucleotides comprise portions of such a polynucleotide sequence having at least about 15 nucleotides and usually at least about 20 nucleotides. Nucleic acid probes comprise portions of such a polynucleotide sequence having fewer nucleotides than about 6 kb, usually fewer than about 1 kb. After appropriate testing to eliminate false positives, these probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250).

The term "probes" includes naturally occurring or recombinant or chemically synthesized single- or double-stranded nucleic acids. They may be labeled by nick translation, Klenow fill-in reaction, PCR or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory,

NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (e.g., hybridization to filter-bound DNA under in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (e.g., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in Example 4.

In instances wherein hybridization of deoxyoligonucleotides is concerned, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligos), 48°C (for 17-base oligos), 55°C (for 20-base oligos), and 60°C (for 23-base oligos).

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., *E. coli*, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an N-terminal methionine residue. This

residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are fragments which induce the expression or an operably linked ORF in response to a specific regulatory factor or physiological event.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below.

The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is

determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the term "biologically active" with reference to angiopoietins means that the polypeptide retains at least one of the biological activities, preferably the activity of one of the human angiopoietins while the term "immunologically active" with reference to angiopoietins means that the polypeptide retains at least one of the immunologic or antigenic activities of one of the human angiopoietins.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as ornithine, which do not normally occur in human proteins.

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, such as angiopoietin activity, may be found by comparing the sequence of the particular polypeptide with that of homologous human or other mammalian angiopoietin peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, i.e., conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity,

hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are typically in the range of about 1 to 5 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 20% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.2 or less). Such a sequence is said to have 80% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 10% (90% sequence identity); in a variation of this embodiment, by no more

than 5% (95% sequence identity); and in a further variation of this embodiment, by no more than 2% (98% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention generally have at least 95% sequence identity with a listed amino acid sequence, whereas substantially equivalent nucleotide sequence of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method.

Nucleic acid sequences encoding such substantially equivalent sequences, e.g., sequences of the recited percent identities, can routinely be isolated and identified via standard hybridization procedures well known to those of skill in the art.

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, often at least about 7 amino acids, typically at least about 9 to 13 amino acids, and, in various embodiments, at least about 17 or more amino acids. To be active, any polypeptide must have sufficient length to display biologic and/or immunologic activity.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change

characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

The term "activated" cells as used in this application are those which are engaged in extracellular or intracellular membrane trafficking, including the export of neurosecretory or enzymatic molecules as part of a normal or disease process.

The term "purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, e.g., polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99.8% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration.

The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed.

The term "intermediate fragment" means a nucleic acid between 5 and 1000 bases in length, and preferably between 10 and 40 bp in length.

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell

in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2): 134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Each of the above terms is meant to encompasses all that is described for each, unless the context dictates otherwise.

NUCLEIC ACIDS AND POLYPEPTIDES OF THE INVENTION

Nucleotide and amino acid sequences of the invention are reported below. Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein-IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form of such protein may be obtained by expression of the disclosed full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where protein of the present invention is membrane bound, soluble forms of the protein are also provided. In

such forms part or all of the regions causing the protein to be membrane bound are deleted so that the protein is fully secreted from the cell in which it is expressed.

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides. The compositions of the present invention include isolated polynucleotides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, novel isolated polypeptides, and antibodies that specifically recognize one or more epitopes present on such polypeptides. Species homologs of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

2. NUCLEIC ACIDS OF THE INVENTION

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide encoding a polypeptide comprising the amino acid sequence of SEQ ID

NO: 2, 4, 6, 8, 10, 12, 46, or 48. A preferred nucleic acid sequence is set forth in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47.

The isolated polynucleotides of the invention further include, but are not limited to a polynucleotide comprising the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47; a polynucleotide comprising the full length protein coding sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47; and a polynucleotide comprising the nucleotide sequence of the mature protein coding sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47. The polynucleotides of the present invention also include, but are not limited to, polynucleotides that encode polypeptides with angiopoietin activity and that hybridize under stringent hybridization conditions to the complement of either (a) the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47, or (b) a nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48; a polynucleotide which is an allelic variant of any polynucleotide recited above; a polynucleotide which encodes a species homolog of any of the proteins recited above; or a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptide of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have at least about 65%, more typically at least about 70%, 75%, 80%, 85% or 90%, and even more typically at least about 95%, sequence identity to a polynucleotide recited above. The invention also provides the complement of the polynucleotides including a nucleotide sequence that has at least about 80%, more typically at least about 90%, and even more typically at least about 95%, sequence identity to a polynucleotide encoding a polypeptide recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions which can routinely isolate polynucleotides of the desired sequence identities.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polypeptides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The sequences falling within the scope of the present invention are not limited to the specific sequences herein described, but also include allelic variations thereof. Allelic variations can be routinely determined by comparing the sequence provided in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47, or a representative fragment thereof, or a nucleotide sequence at least 99.9% identical to SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47 with a sequence from another isolate of the same species.

To accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another which encodes the same amino acid is expressly contemplated. Any specific sequence disclosed herein can be readily screened for errors by resequencing a particular fragment, such as an ORF, in both directions (i.e., sequence both strands).

The present invention further provides recombinant constructs comprising a nucleic acid having the sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having the sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47 or a fragment thereof is inserted, in a forward or reverse orientation.

In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. For vectors comprising the EMFs and UMFs of the present invention, the vector may further comprise a marker sequence or heterologous ORF operably linked to the EMF or UMF. Large numbers of suitable vectors and promoters are known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P_R, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream

structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli, Bacillus subtilis, Salmonella typhimurium* and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequences that hybridize under stringent conditions to a fragment of the DNA

sequence of SEQ ID NO: 1, which fragment is greater than about 10 bp, preferably 20-50 bp, and even greater than 100 bp. In accordance with the invention, polynucleotide sequences which encode the novel nucleic acids, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. The amino acid sequence variants of the nucleic acids are preferably constructed by mutating the polynucleotide to give an amino acid sequence that does not occur in nature. These amino acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells, and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein..

In a preferred method, polynucleotides encoding the novel nucleic acids are changed via site-directed mutagenesis. This method uses oligonucleotide sequences that encode the polynucleotide sequence of the desired amino acid variant, as well as a sufficient adjacent nucleotide on both sides of the changed amino acid to form a stable

duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., DNA 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives the desired amino acid variant.

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., Gene 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and Current Protocols in Molecular Biology, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

3. HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of angiopoietin DNA sequences allows for modification of cells to permit, or increase, expression of endogenous angiopoietins. Cells can be modified (*e.g.*, by homologous recombination) to provide increased angiopoietin expression by replacing, in whole or in part, the naturally occurring angiopoietin promoter with all or part of a heterologous promoter so that the cells express angiopoietin at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to angiopoietin encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the angiopoietin coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the angiopoietin coding sequences in the cells.

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., Basic Methods in Molecular Biology (1986)). The host cells containing one of polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs

of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell tines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast, insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains,

Candida, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include Escherichia coli, Bacillus subtilis, Salmonella typhimurium, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting, including polyadenylation signals. mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable

marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

Exemplary gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4. POLYPEPTIDES OF THE INVENTION

The polynucleotides of SEQ ID NO. 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47 encode the angiopoietin polypeptide sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48. The polypeptide set out in SEQ ID NO. 2, 4, 6, 8, 10, 12, 46, and 48 display amino acid homology with human angiopoietins Ang-1, Ang-2, Ang-4, Ang-Y and the human angiopoietin-like protein, and thus represent novel molecules within the angiopoietin family. Additional angiopoietin family members can be identified using SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47 as a molecular probe.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising the amino acid sequence of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 46, or 48 or the amino acid sequence encoded by the DNA of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47 or a portion thereof corresponding to the full length or mature protein. Polypeptides of the invention also include polypeptides with angiopoietin activity that are encoded by (a) the polynucleotide of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47, or (b)

polynucleotides encoding SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48 or (b) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. Biologically active or immunologically active variants of the angiopoietin protein sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48 and "substantial equivalents" thereof (e.g., with 65%, 70%, 75%, 80%, 85%, 90%, typically 95%, more typically 98% or most typically 99% amino acid identity) that retain angiopoietin activity, preferably human angiopoietin activity, are also contemplated. Polypeptides encoded by allelic variants may have a similar or increased or decreased activity compared to the polypeptides of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The invention also relates to methods for producing a polypeptide comprising growing a culture of the cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the cells or the culture medium, and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins. A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. This is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating

antibodies against the native polypeptide. In an alternative method, the polypeptide or protein is purified from host cells which produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that retain biological/immunological activity include fragments encoding greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains. For example, preferred fibrinogen polypeptide fragments of the invention comprise amino acid residues 292 to 329, residues 333 to 346, and 450 to 475 in SEQ ID NO: 2 (CG006alt2), residues 25 to 38 and 142 to 167 in SEQ ID NO: 4 (CG006alt3); residues 258 to 295, residues 386 to 415, residues 420 to 445, and residues 299 to 312 in SEQ ID NO: 6 (CG007); residues 219 to 248, residues 252 to 277, 135 to 148, residues 200 to 214, and residues 182 to 200 in SEQ ID NO: 46 (CG015alt1); residues 157 to 194, 282 to 311, 315 to 340, 198 to 211, 263 to 277 and 245 to 263 in SEQ ID NO: 48 (CG015alt2); and residues 193 to 230, residues 337 to 366, residues 307 to 321, residues 283 to 301, residues 234 to 247 in SEQ ID NO: 8 (CG144).

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention. The purified polypeptides can be used in in vitro binding assays which are well known in the art to identify molecules which bind to the polypeptides.

There are a number of different libraries used for the identification of small molecules that bind a polypeptide of the invention, including, (1) chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

Chemical libraries consist of structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening. Natural product libraries are collections of microorganisms, animals, plants, or marine organisms which are used to create mixtures for screening by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of plants or marine organisms. Natural product libraries include polyketides, non-ribosomal peptides, and variants (non-naturally occurring) variants thereof. For a review, see Science 282:63-68 (1998). Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds as a mixture. They are relatively easy to prepare by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in in vivo tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the binding molecules may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the polypeptide

of the invention or binding molecules may be complexed with imaging agents for targeting and imaging, e.g., areas of vascularization.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBat.RTM. kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl.RTM. or Cibacrom blue 3GA Sepharose.RTM.; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and In Vitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The

protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include angiopoietin analogs or variants. This embraces fragments of angiopoietin of the invention, as well as analogs (variants) of angiopoietin in which one or more amino acids has been deleted, inserted, or substituted. Analogs of the invention also embrace fusions or modifications of angiopoietin wherein the angiopoietin or analog is fused to another moiety or moieties, e.g., targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to angiopoietin or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to desired cell types. Other moieties which may be fused to angiopoietin or an analog include therapeutic agents which are used for treatment of indications as described herein.

5. GENE THERAPY

Mutations in the angiopoietin gene that result in loss of normal function of the angiopoietin comprehends gene therapy to restore angiopoietin activity would thus be indicated in treating those disease states (for example, various forms of cancer described herein). Delivery of a functional angiopoietin gene to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (*e.g.*, adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of angiopoietin will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of angiopoietin.

5.1 TRANSGENIC ANIMALS

In methods to determine biological functions of angiopoietins in vivo, one or more angiopoietin genes are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the role(s) angiopoietins play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate angiopoietin activity. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of an angiopoietin promoter is either activated or inactivated to alter the level of expression of the angiopoietin protein. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous angiopoietin promoter to provide for increased angiopoietin expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

6. USES AND BIOLOGICAL ACTIVITY

The angiopoietin activity of a polypeptide of the invention may manifest as, e.g., anti-angiogenic activity or angiogenesis promoting activity. The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for

introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether angiopoietin polypeptides, binding partners thereof, or inhibitors thereof would be beneficial to the subject in need of treatment.

Angiogenesis plays a role in chronic inflammation, including chronic pancreatitis, dermatosis associated with chronic inflammation, including psoriasis, cirrhosis, asthma, multiple sclerosis, arthritis, including rheumatoid arthritis, reactive arthritis and chronic inflammatory arthritis, autoimmune disorders, including vasculitis, glomerulonephritis, experimental allergic encephalomyelitis (EAE), lupus, myasthenia gravis, ulcerative colitis, Crohn's disease, inflammatory bowel disease, chronic inflammation associated with hemodialysis, granulocyte transfusion associated syndrome; rejection reactions after allograft and xenograft transplantation, including graft versus host disease; and other chronic inflammatory disorders.

Angiogenesis in the eye is involved in ocular neovascularization, proliferative retinopathy, macular degeneration, and diabetic ocular disease, in particular, diabetic iris neovascularization and retinopathy.

Coronary atheroma are highly vascularized by a fragile capillary network, and rupture of these newly formed capillaries when they are exposed to high intravascular pressures may lead to hemorrhage into atherosclerotic plaques and vessel occlusion. Inhibition of angiogenesis thus may reduce the growth of atherosclerotic plaques and may be useful in the treatment of atherosclerosis, ischemic heart disease, myocardial infarction, coronary heart disease, restenosis, particularly following balloon angiography, neointimal hyperplasia, disruption of intercellular junctions in vascular endothelium, hypertension, vessel injury, arterial ischemia, arterial stenosis, peripheral vascular disease, stroke, coronary obstruction, and periventricular leukomalacia, chronic cor pulmonalea (disease of the right or both ventricle(s) of the heart), and other conditions associated with decreased or increased myocardial revascularization. New angiopoietin family members are also expected to be useful in vascular remodeling as an alternative to coronary artery bypass surgery to prevent myocardial infarction, and useful to recruit progenitor cells into the hematopoietic lineage to treat anemia.

Introduction of angiogenic factors into ischemic myocardium is expected to enhance the development of collateral vessels, accelerate healing of necrotic tissue, and

prevent infarct expansion and cardiac dilation. Similarly, essential hypertension is based on an impaired capacity for vascular growth.

Methods of the invention also include treatment for cardiovascular conditions and pathologies including modified microvascular hyperpermeability, hemostasis, microvascular disease associated with impaired angiogenesis, pulmonary vascular disorders in portal hypertension, and capillary leak syndrome. New angiopoietin family members are also expected to be useful in enhancing the strength and integrity of vessels, possibly decreasing the likelihood of vessel rupture and associated artery blockage at sites of atherosclerotic plaques. Polypeptides of the invention will also be useful in treating causes of thrombolytic disease or thrombocytopaenia.

In addition, angiopoietin family members are expected to be used to treat stem cells in vivo, in vitro or ex vivo to produce hemangioblasts to augment these cell types in a variety of human pathologies or for research into the function or development of these cells.

Angiogenesis is also important in bone conditions including osteoporosis, osteoradionecrosis, osteonecrosis generally, osteonecrosis of the femoral head, fracture healing and repair generally, fracture healing associated with autogenous and allogeneic bone grafts, and necrosis and hypoxia of bone adjacent a fracture.

Angiogenesis also occurs during the female reproductive cycle and is involved in endometriosis, uterine fibroids, other conditions associated with dysfunctional vascular proliferation (including endometrial microvascular growth) during the female reproductive cycle.

Angiogenesis is also involved in abnormal vascular growth, including cerebral arteriovenous malformations (AVMs), gastrointestinal mucosal injury and repair, ulceration of the gastroduodenal mucosa in patients with a history of peptic ulcer disease, including ischemic tissue resulting from stroke, a wide spectrum of pulmonary vascular disorders in liver disease and portal hypertension in patients with nonhepatic portal hypertension, including hepatopulmonary syndrome and pulmonary hypertension (portopulmonary hypertension), hemangiopericytoma, pyogenic granuloma, liver failure, and autoimmune diseases.

Angiogenesis is also of considerable importance in cancer conditions because new vessel production is required to support the rapid growth of cancer cells. Inhibition of angiogenesis thus may promote tumor regression in adult and pediatric oncology, including reducing growth of solid tumors/malignancies, locally advanced tumors, metastatic cancer, human soft tissue sarcomas, cancer metastases, including lymphatic metastases, blood cell malignancies, including multiple myeloma, leukemias, effusion lymphomas (body cavity based lymphomas), lung cancer, including small cell carcinoma, non-small cell cancers, breast cancer, including small cell carcinoma and ductal carcinoma, gastrointestinal cancers, including stomach cancer, colon cancer, colorectal cancer, polyps associated with colorectal neoplasia, pancreatic cancer, liver cancer, urological cancers, including bladder cancer, prostate cancer, malignancies of the female genital tract, including ovarian carcinoma, uterine endometrial cancers, and solid tumors in the ovarian follicle, kidney cancer, including renal cell carcinoma, brain cancer, including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers, including osteomas, skin cancers, including malignant melanoma, tumor progression of human skin keratinocytes, and squamous cell cancer, hemangiopericytoma, and Kaposi's sarcoma.

Additional uses for polypeptides of the present invention, as well as modulators there of, are described below.

Polypeptides of the invention may also possess one or more tenascin-like biological activities. Tenascins are extracellular matrix proteins involved in regulation of developmental processes, such as morphogenetic cell migration and organogenesis of many organs and tissues, as indicated by tissue distribution and regulated expression during embryogenesis. Known members of the gene family include tenascin/cytotactin (tenascin-C), restrictin/J-160/180 (tenascin-R), and the tenascin-like gene present in the major histocompatibility complex class III locus (tenascin-X). The tenascins are multimeric extracellular matrix glycoproteins with multiple isoforms arising from alternative splicing. The proteins have repeated structural domains, including heptad repeats, epidermal growth factor (EGF)-like repeats, fibronectin type III repeats, and globular domains found in fibrinogens. Tenascin-R appears to be expressed specifically in the central and peripheral nervous system, tenascin-X is most prominently expressed in

skeletal and heart muscle, while tenascin-C is most highly expressed in many developing tissues, including the nervous system, but not in skeletal and heart muscle. Overexpression of tenascin-C is also observed in tumors.

Previous reports suggest that tenascin-C is an adhesion-modulating protein. The protein is highly conserved across species boundaries and is expressed in a variety of tissues. In the nervous systems of rodents and chickens, tenascin-C is predominantly expressed at early developmental ages and may be involved in different steps of neural development. Tenascin-C activity has been implicated in synaptogenesis, migration of different neural cell types, axonal growth in the developing and lesioned nervous system, and in the formation and maintenance of discrete anatomical boundaries. Experiments with anti-tenascin-C monoclonal antibodies and tenascin-C polypeptide fragments suggest different functions of tenascin-C are associated with different domains of the protein. Most cells do not express tenascin-C constitutively, but expression is induced by growth factors and hormones, such as transforming growth factor β and interleukin-1. Even though tenascin-C has anti-adhesive properties, the protein appears to influence the differentiation of a variety of cell types. Expression of tenascin-C in tumors has led to development of radio-labeled monoclonal anti-tenascin-C antibodies for targeting tumor therapy.

Modulation of tenascin activity can be useful in many pathological conditions, including pre-eclampsia decidua, neurodegeneration, abnormal embryonic development, abnormal wound healing, conditions associated with neoplastic growth, large-bowel diseases generally and specifically ulcerative colitis, small axillary node-negative breast carcinomas and distant metastasis, colorectal carcinomas, inflammation in general, chronic and seasonal asthma, abnormal osteoblastic differentiation, tendon disease including abnormal tendon formation and degenerate tendons, abnormal collagen fibril organization, mononuclear cell infiltration, angiopoiesis, chondrogenic tumors, proliferative activity of tumor cells in enchondromas and chondrosarcomas, alterations of extracellular matrix, tumor development, active scar formation, granulomas in sarcoidosis, cryptic fibrosing alveolitis (CFA), abnormal assembly and activity of focal adhesions, neointima formation after acute vascular injury, new growth and expansion

within primary atherosclerotic plaques, and intimal repair and luminal narrowing in restenosis after angioplasty.

6.1. RESEARCH USES AND UTILITIES

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially

binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

6.2. NUTRITIONAL USES

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

6.3. CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations.

6.4. IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. For example, polypeptides of the invention may be used to modulate the immune response in the treatment of leukopaenia, immune coagulation, inflammatory reactions and autoimmune disease.

6.5. HEMATOPOIESIS REGULATING ACTIVITY

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis.

6.6. TISSUE GROWTH ACTIVITY

A protein of the present invention, particularly proteins that promote angiogenesis or vascularization, also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers, and in treatment of conditions involving hypovascularization.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also

be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, i.e. for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's

disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, , gastric ulcers, surgical and traumatic wounds, burns and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.),

Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

6.7. CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes.

6.8. HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A protein of the invention may also exhibit hemostatic or thrombolytic activity. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

6.10. RECEPTOR/LIGAND ACTIVITY

A protein of the present invention may also demonstrate activity as receptors , receptor ligands or inhibitors or agonists of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics.

By way of example, the angiopoietin polypeptides of the invention may be used as a ligand for a receptor thereby modulating (*i.e.*, enhancing or inhibiting) the biological activity of that receptor. Whether the angiopoietin polypeptides of the invention exhibit agonist, partial agonist, antagonist, or partial antagonist activity for a particular receptor, such as a Tie-2 receptor, in a particular cell type can be determined by conventional techniques known to those skilled in the art, such as those described below in sections 6.11.1 and 6.11.2 and in the Examples below. Examples of cells that may be contacted with the protein of the invention include, but are not limited to, mammalian cells such as

endothelial cells. Preferably the novel protein of the invention acts as an antagonist for a Tie-2 receptor so that the biological activities of that receptor are inhibited.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention are expected to exhibit an affinity for Tie-2. Thus, the polypeptides of the present invention may be used, for example, as competitors in assays involving Tie-2. Alternatively, the polypeptides of the invention may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego) and used in both in vivo and in vitro to bind to Tie-2. Examples of radioisotopes include, but are not limited to, tritium and carbon-14 . Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin. By way of example, the proteins coupled to such molecules are useful in studies involving in vivo or in vitro metabolism of angiopoietin.

6.11 DRUG SCREENING WITH ANGIOPOIETIN POLYPEPTIDES

This invention is particularly useful for screening compounds by using the angiopoietin polypeptides of the invention, particularly binding fragments, in any of a variety of drug screening techniques. The polypeptides employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the desired angiopoietin polypeptide. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between angiopoietin polypeptides of the invention and the agent being tested or examine the diminution in complex formation between the angiopoietin polypeptides and an appropriate cell line, which are well known in the art.

6.11.1 ASSAY FOR RECEPTOR BINDING ACTIVITY

The invention also provides methods to detect specific binding of an angiopoietin polypeptide of the invention to a binding partner polypeptide, and in particular a receptor polypeptide. Receptors expected to be useful in binding assays of this type include Tie-2, Tie-1, and other binding partner/receptors identified using assay well known and routinely practiced in the art.

In one embodiment, receptor antagonist activity of the angiopoietin polypeptides of the invention is determined using a method that involve (1) forming a mixture comprising angiopoietin, receptor, and/or its agonists and antagonists (or agonist or antagonist drug candidates) and/or antibodies specific for the angiopoietin polypeptides of the invention; (2) incubating the mixture under conditions whereby, but for the presence of said angiopoietin polypeptide of the invention and/or agonists and antagonists (or agonist or antagonist drug candidates) and/or antibodies specific for the angiopoietin polypeptides of the invention, the angiopoietin binds to the receptor; and (3) detecting the presence or absence of specific binding of angiopoietin to the receptor.

The art provides numerous assays particularly useful for identifying previously unknown binding partners for angiopoietins of the invention. For example, expression cloning, using mammalian or bacterial cells, can be used to identify polynucleotides encoding angiopoietin binding partners. As another example, affinity chromatography with an immobilized angiopoietin polypeptide can be used to isolate polypeptides that recognized and bind an angiopoietin of the invention. As still another example, overlay assays can be used to identify binding partner polypeptides.

6.11.2 ASSAY FOR ANTAGONISTS AND AGONISTS OF ANGIOPOIETIN RECEPTOR ACTIVITY

Numerous techniques are known in the art to assay for agonists and antagonists of angiopoietin receptor activity. For example, the mouse cornea (micropocket) neovascularization assay [Asahara, et al., Circ. Res 83:233-240 (1998)] permits in vivo analysis of both agonists and antagonists activities for angiopoietins.

Other assays previously described include determination of receptor phosphorylation following angiopoietin binding in endothelial cells and fibroblasts

expressing an angiopoietin receptor [Davis, et al., Cell 87:1161-1169 (1996); Maisonpierre, et al., Science 277:48-50 (1997)].

In still another assay, vessel formation is measured as described in Koblizek, et al., Curr. Biol. 8:529-532 (1998). Assays can be performed with or without competitive inhibitors of angiopoietin receptor binding, such as monoclonal antibodies and/or Ang-2.

As another example, angiogenesis can be assessed using the Matrigel™ model as previously described [Passaniti, et al., Lab. Invest. 67:519-528 (1992)]. This model uses a Matrigel™ basement membrane preparation mixed with FGF-2 and heparin, which induces intense neovascularization within the gel when injected subcutaneously into mice. The extent of angiogenesis is quantitated by measuring the hemoglobin content of the gels. Compounds that neutralize the angiogenic properties of heparin will inhibit angiogenesis in the model.

6.12. ANTI-INFLAMMATORY ACTIVITY

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. In particular, the angiopoietin polypeptides of this invention may be utilized to prevent or treat condition such as, but not limited to, utilized, for example, as part of methods for the prevention and/or treatment of disorders involving sepsis, acute

pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

6.13. LEUKEMIAS

Leukemias and related disorders may be treated or prevented by administration of a therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monotypic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

6.14. NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

(i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;

(ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

(iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;

(iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;

(v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;

(vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;

(vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and

(viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

(ii) increased sprouting of neurons in culture or in vivo;

(iii) increased production of a neuron-associated molecule in culture or in vivo, e.g., choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or

(iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In a specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

6.15. OTHER ACTIVITIES

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or

caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

6.16 IDENTIFICATION OF POLYMORPHISMS

The demonstration of polymorphisms, for example the polymorphisms illustrated below, makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving vascular stability or neovascularization) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to neovascularization makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the

DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or absence of the polymorphism) may be performed.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

7. THERAPEUTIC METHODS

The novel angiopoietin polypeptides (including fragments, analogs and variants) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified below.

7.1 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines,

growth factors, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, GM-CSF, thrombopoietin, stem cell factor, and erythropoietin. Particularly preferred are compositions that include other known angiopoietins, for example Ang-1, Ang-2, Ang-4, Ang-Y, and/or the human angiopoietin-like polypeptide, and/or vascular endothelial growth factor (VEGF). Preferred growth factors for use in pharmaceutical compositions of the invention include angiogenin, bone morphogenic protein-1, bone morphogenic protein-2, bone morphogenic protein-3, bone morphogenic protein-4, bone morphogenic protein-5, bone morphogenic protein-6, bone morphogenic protein-7, bone morphogenic protein-8, bone morphogenic protein-9, bone morphogenic protein-10, bone morphogenic protein-11, bone morphogenic protein-12, bone morphogenic protein-13, bone morphogenic protein-14, bone morphogenic protein-15, bone morphogenic protein receptor IA, bone morphogenic protein receptor IB, brain derived neurotrophic factor, ciliary neutrophic factor, ciliary neutrophic factor receptor α
cytokine-induced neutrophil chemotactic factor 1, cytokine-induced neutrophil, chemotactic factor 2 α, cytokine-induced neutrophil chemotactic factor 2 β, β endothelial cell growth factor, endothelin 1, epidermal growth factor, epithelial-derived neutrophil attractant, fibroblast growth factor 4, fibroblast growth factor 5, fibroblast growth factor 6 fibroblast growth factor 7, fibroblast growth factor 8, fibroblast growth factor 8b, fibroblast growth factor 8c, fibroblast growth factor 9, fibroblast growth factor 10, fibroblast growth factor acidic, fibroblast growth factor basic, glial cell line-derived neutrophic factor receptor α 1, glial cell line-derived neutrophic factor receptor α 2, growth related protein, growth related protein α, growth related protein β, growth related protein γ, heparin binding epidermal growth factor, hepatocyte growth factor, hepatocyte growth factor receptor, insulin-like growth factor I, insulin-like growth factor receptor, insulin-like growth factor II, insulin-like growth factor binding protein, keratinocyte growth factor, leukemia inhibitory factor, leukemia inhibitory factor receptor α, nerve growth factor nerve growth factor receptor, neurotrophin-3, neurotrophin-4, placenta growth factor, placenta growth factor 2, platelet-derived endothelial cell growth factor, platelet derived growth factor, platelet derived growth factor A chain, platelet derived

growth factor AA, platelet derived growth factor AB, platelet derived growth factor B chain, platelet derived growth factor BB, platelet derived growth factor receptor α, platelet derived growth factor receptor β, pre-B cell growth stimulating factor, stem cell factor, stem cell factor receptor, transforming growth factor α, transforming growth factor β, transforming growth factor β1, transforming growth factor β1.2, transforming growth factor β2, transforming growth factor β3, transforming growth factor β5, latent transforming growth factor β1, transforming growth factor β binding protein I, transforming growth factor β binding protein II, transforming growth factor β binding protein III, tumor necrosis factor receptor type I, tumor necrosis factor receptor type II, urokinase-type plasminogen activator receptor, vascular endothelial growth factor, and chimeric proteins and biologically or immunologically active fragments thereof

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent. A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein.

Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in

rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

7.2. ROUTES OF ADMINISTRATION

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the

scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

7.3. COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution.

The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to

the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions

may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a cosolvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The cosolvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or

carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the compounds of the invention may be provided as salts with pharmaceutically compatible counterions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the

invention. The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent. Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 μg to about 100 mg (preferably about 0.1 μg to about 10 mg, more preferably about 0.1 μg to about 1 mg) of protein of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon, ligament, or other tissue regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention.

Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which

represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the fracture repair activity of the progenitor cells. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-.alpha. and TGF-.beta.), insulin-like growth factor (IGF), other known angiopoietins, VEGF, bone morphogenic protein (BMP), as well as other cytokines and/or growth factors described herein.

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

7.4. EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of

the subject being treated. Determination of the effective amounts is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture (i.e., the concentration of the test compound which achieves a half-maximal inhibition of the C-proteinase activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD_{50} and ED_{50}. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the C-proteinase inhibiting effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from in vitro data; for example, the concentration necessary to achieve 50-90% inhibition of the C-proteinase using the assays described herein. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for the human angiopoietin polypeptides of the invention will be in the range of about 0.01 to 100 mg/kg of body weight daily, with the preferred dose being about 0.1 to 25 mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

7.5. PACKAGING

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

8. ANTIBODIES

Another aspect of the invention is an antibody that specifically binds the polypeptide of the invention. Such antibodies include monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, bifunctional/bispecific antibodies, humanized antibodies, human antibodies, and complementary determining region (CDR)-grafted antibodies, including compounds which include CDR and/or antigen-binding sequences, which specifically recognize a polypeptide of the invention. Preferred antibodies of the invention are human antibodies which are produced and

identified according to methods described in WO93/11236, published June 20, 1993, which is incorporated herein by reference in its entirety. Antibody fragments, including Fab, Fab', $F(ab')_2$, and F_v, are also provided by the invention. The term "specific for" indicates that the variable regions of the antibodies of the invention recognize and bind angiopoietin polypeptides exclusively (*i.e.*, able to distinguish an angiopoietin polypeptides from the family of angiopoietin polypeptides despite sequence identity, homology, or similarity found in the family of polypeptides), but may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow et al. (Eds), Antibodies A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor , NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the angiopoietin polypeptides of the invention are also contemplated, provided that the antibodies are first and foremost specific for, as defined above, angiopoietin polypeptides. As with antibodies that are specific for full length angiopoietin polypeptides, antibodies of the invention that recognize angiopoietin fragments are those which can distinguish angiopoietin polypeptides from the family of angiopoietin polypeptides despite inherent sequence identity, homology, or similarity found in the family of proteins. Antibodies of the invention can be produced using any method well known and routinely practiced in the art.

Non-human antibodies may be humanized by any methods known in the art. In one method, the non-human CDRs are inserted into a human antibody or consensus antibody framework sequence. Further changes can then be introduced into the antibody framework to modulate affinity or immunogenicity.

Antibodies of the invention are useful for, for example, therapeutic purposes (by modulating activity of a polypeptide of the invention), diagnostic purposes to detect or quantitate a polypeptide of the invention, as well as purification of a polypeptide of the invention. Kits comprising an antibody of the invention for any of the purposes described herein are also comprehended. In general, a kit of the invention also includes a

control antigen for which the antibody is immunospecific. The invention further provides a hybridoma that produces an antibody according to the invention. Antibodies of the invention are useful for detection and/or purification of the polypeptides of the invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R. P. Merrifield, J. Amer. Chem. Soc. 85, 2149-2154 (1963); J. L. Krstenansky, et al., FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein. In general, techniques for preparing polyclonal and monoclonal antibodies as well as hybridomas capable of producing the desired antibody are well known in the art (Campbell, A.M., Monoclonal Antibodies Technology: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1984); St. Groth et al., J. Immunol. 35:1-21 (1990); Kohler and Milstein, Nature 256:495-497 (1975)), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., Immunology Today 4:72 (1983); Cole et al., in Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc. (1985), pp. 77-96).

Any animal (mouse, rabbit, etc.) which is known to produce antibodies can be immunized with a peptide or polypeptide of the invention. Methods for immunization are well known in the art. Such methods include subcutaneous or intraperitoneal injection of the polypeptide. One skilled in the art will recognize that the amount of the protein encoded by the ORF of the present invention used for immunization will vary based on the animal which is immunized, the antigenicity of the peptide and the site of injection.

The protein that is used as an immunogen may be modified or administered in an adjuvant in order to increase the protein's antigenicity. Methods of increasing the antigenicity of a protein are well known in the art and include, but are not limited to, coupling the antigen with a heterologous protein (such as globulin or β-galactosidase) or through the inclusion of an adjuvant during immunization.

For monoclonal antibodies, spleen cells from the immunized animals are removed, fused with myeloma cells, such as SP2/0-Ag14 myeloma cells, and allowed to become monoclonal antibody producing hybridoma cells. Any one of a number of methods well known in the art can be used to identify the hybridoma cell which produces an antibody with the desired characteristics. These include screening the hybridomas with an ELISA assay, western blot analysis, or radioimmunoassay (Lutz et al., Exp. Cell Research. 175:109-124 (1988)). Hybridomas secreting the desired antibodies are cloned and the class and subclass is determined using procedures known in the art (Campbell, A.M., Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1984)). Techniques described for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce single chain antibodies to proteins of the present invention.

For polyclonal antibodies, antibody containing antiserum is isolated from the immunized animal and is screened for the presence of antibodies with the desired specificity using one of the above-described procedures. The present invention further provides the above-described antibodies in delectably labeled form. Antibodies can be delectably labeled through the use of radioisotopes, affinity labels (such as biotin, avidin, etc.), enzymatic labels (such as horseradish peroxidase, alkaline phosphatase, etc.) fluorescent labels (such as FITC or rhodamine, etc.), paramagnetic atoms, etc. Procedures for accomplishing such labeling are well-known in the art, for example, see (Sternberger, L.A. et al., J. Histochem. Cytochem. 18:315 (1970); Bayer, E.A. et al., Meth. Enzym. 62:308 (1979); Engval, E. et al., Immunol. 109:129 (1972); Goding, J.W. J. Immunol. Meth. 13:215 (1976)).

The labeled antibodies of the present invention can be used for in vitro, in vivo, and in situ assays to identify cells or tissues in which a fragment of the polypeptide of

interest is expressed. The antibodies may also be used directly in therapies or other diagnostics. The present invention further provides the above-described antibodies immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and Sepharose™, acrylic resins and such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir, D.M. et al., "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10 (1986); Jacoby, W.D. et al., Meth. Enzym. 34 Academic Press, N.Y. (1974)). The immobilized antibodies of the present invention can be used for in vitro, in vivo, and in situ assays as well as for immuno-affinity purification of the proteins of the present invention.

9. COMPUTER READABLE SEQUENCES

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs

and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention. By providing the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47 or a representative fragment thereof, or a nucleotide sequence at least 99.9% identical to SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means

which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

10. TRIPLE HELIX FORMATION

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA. Polynucleotides suitable for use in these methods are usually 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix- formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

11. DIAGNOSTIC ASSAYS AND KITS

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the

polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample. In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

12. MEDICAL IMAGING

The novel angiopoietin polypeptides of the invention are useful in medical imaging, e.g., imaging the site of neovascularization and other sites having Tie-2 receptor antagonist receptor molecules. See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled angiopoietin polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled angiopoietin polypeptide in vivo at the target site.

13. SCREENING ASSAYS

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by the ORF from a polynucleotide of the invention to a specific domain of the polypeptide encoded by a polypeptide of the invention. In detail, said method comprises the steps of:

(a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and

(b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex. so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be

selected and screened at random or rationally selected or designed using protein modeling techniques.

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like capable of binding to a specific peptide sequence in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods usually contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix- formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been

demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents. Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent, in the control of bacterial infection by modulating the activity of the protein encoded by the ORF. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

14. USE OF NUCLEIC ACIDS AS PROBES

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from the nucleotide sequence of the SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47. Because the corresponding gene is only expressed in a limited number of tissues, especially adult tissues, a hybridization probe derived from SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described US Patent Nos 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for

mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals. The nucleotide sequence may be used to produce purified polypeptides using well known methods of recombinant DNA technology. Among the many publications that teach methods for the expression of genes after they have been isolated is Goeddel (1990) Gene Expression Technology, Methods and Enzymology, Vol 185, Academic Press, San Diego. Polypeptides may be expressed in a variety of host cells, either prokaryotic or eukaryotic. Host cells may be from the same species from which a particular polypeptide nucleotide sequence was isolated or from a different species. Advantages of producing polypeptides by recombinant DNA technology include obtaining adequate amounts of the protein for purification and the availability of simplified purification procedures.

Each sequence so obtained was compared to sequences in GenBank using a search algorithm developed by Applied Biosystems and incorporated into the INHERIT™ 670 Sequence Analysis System. In this algorithm, Pattern Specification Language (developed by TRW Inc., Los Angeles, CA) was used to determine regions of homology. The three parameters that determine how the sequence comparisons run were window size, window offset, and error tolerance. Using a combination of these three parameters, the DNA

database was searched for sequences containing regions of homology to the query sequence, and the appropriate sequences were scored with an initial value. Subsequently, these homologous regions were examined using dot matrix homology plots to distinguish regions of homology from chance matches. Smith-Waterman alignments were used to display the results of the homology search. Peptide and protein sequence homologies were ascertained using the INHERIT™ 670 Sequence Analysis System in a way similar to that used in DNA sequence homologies. Pattern Specification Language and parameter windows were used to search protein databases for sequences containing regions of homology that were scored with an initial value. Dot-matrix homology plots were examined to distinguish regions of significant homology from chance matches.

Alternatively, BLAST, which stands for Basic Local Alignment Search Tool, is used to search for local sequence alignments (Altschul SF (1993) J Mol Evol 36:290-300; Altschul, SF et al (1990) J Mol Biol 215:403-10). BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologs. Whereas it is ideal for matches which do not contain gaps, it is inappropriate for performing motif-style searching. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP). An HSP consists of two sequence fragments of arbitrary but equal lengths whose alignment is locally maximal and for which the alignment score meets or exceeds a threshold or cutoff score set by the user. The BLAST approach is to look for HSPs between a query sequence and a database sequence, to evaluate the statistical significance of any matches found, and to report only those matches which satisfy the user-selected threshold of significance. The parameter E establishes the statistically significant threshold for reporting database sequence matches. E is interpreted as the upper bound of the expected frequency of chance occurrence of an HSP (or set of HSPs) within the context of the entire database search. Any database sequence whose match satisfies E is reported in the program output.

In addition, BLAST analysis was used to search for related molecules within the libraries of the LIFESEQ™ database. This process, an "electronic northern" analysis is analogous to northern blot analysis in that it uses one cellubrevin sequence at a time to search for identical or homologous molecules at a set stringency. The stringency of the

electronic northern is based on "product score". The product score is defined as (% nucleotide or amino acid [between the query and reference sequences] in Blast multiplied by the % maximum possible BLAST score [based on the lengths of query and reference sequences]) divided by 100. At a product score of 40, the match will be exact within a 1-2% error; and at 70, the match will be exact. Homologous or related molecules can be identified by selecting those which show product scores between approximately 15 and 30.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples.

EXAMPLE 1

Cloning of Angiopoietin cDNAs

Novel nucleic acids were obtained from various cDNA libraries (prepared from human mRNA purchased from Invitrogen, San Diego, CA) using standard PCR, sequencing by hybridization (SBH) sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for pSport1 (GIBCO BRL, Grand Island, NY) vector sequences which flank the inserts. These samples were spotted onto nylon membranes and hybridized with oligonucleotide probes to give sequence signatures. The clones were clustered into groups of similar or identical sequences, and single representative clones were selected from each group for gel sequencing. The 5' sequence of the amplified inserts was then deduced using the reverse M13 sequencing primer in a typical Sanger sequencing protocol. PCR products were purified and subjected to flourescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems (ABI) sequencer.

Sequence analysis identified seven polynucleotides encoding five novel polypeptides designated CG006, CG007, CG015, CG144, and CG250. The 5′ sequences for clones CG006, CG007, CG015, and CG144 were determined as described in Example 2. Identification of the various gene sequences was as described below.

Clone CG006

CG006 was identified by combining sequences CG006alt2 (internal designation AngPO1alt2_Hy040999, SEQ ID NO: 1) and CG006alt3 (internal designation AngPO1alt3_Hy060399, SEQ ID NO: 3). The contig encoding CG006 alt2 was deduced from two clones identified in an adult kidney library (clone 2462967, RTA00001120.g.22, and clone 2385177, RTA00002051.i.02) and eight clones in an adult liver library (clone 2850275, RTA00001140.j.11/RTA00003679.b.09/RTA00003679.c.06; clone 2851655, RTA00001804.b.10; clone 2851986, RTA00001804.d.15; clone 2851847, RTA00003679.b.11; clone 2854934, RTA00003679.b.12/ RTA00003679.c.10; clone 2894719, RTA00003679.b.16; clone 2918501, RTA00003679.b.24/RTA00003679.c.01; and clone 3082131, RTA00003679.c.033).

The contig encoding CG006alt3 was deduced from singles clones isolated from adult liver and a first fetal liver/spleen libraries, and seven additional clones from a second fetal liver library (adult liver library clone 2853792, RTA00001804F.o.12; fetal liver/spleen library clone 4963222, RTA00001171F.a.21/RTA00002115F.b.08/RTA00002115F.c.12; clone 17216212, RTA00003682F.a.11; clone 17280808, RTA00003682F.a.12; clone 17447158, RTA00003682F.a.15; clone 17399459, RTA00003682F.a.16; clone 17449878, RTA00003682F.a.17; clone 17121000, RTA00003682R.a.10; and clone 18730012, RTA00003682R.a.18). A contig comprising the 5′ terminus of CG006alt2 and the entire sequence for CG006alt3 is set out in SEQ ID NO: 11.

Clone CG007

CG007, encoded by sequence CG007alt1 (internal designation AngPO2_Hy040299, SEQ ID NO: 5). The contig encoding CG007alt1 was deduced from numerous library clones as set out in Table 1 below.

Table 1
Source of ESTs in CG007alt1 Contig
Library Number Clones (Tissue Distribution)

ABT004 1	ALV002 4	FLG003 3	IB2002 2
ADP001 5	AOV001 3	FLS001 2	IBM002 2
AHR001 2	APL001 2	FLS002 4	IBS001 1
AKT002 32	ATS001 1	FLV001 1	LGT002 3
ALG001 1	BLD001 1	FSK001 1	MMG0018
ALV001 3	CVX001 1	FUC001 1	PRT001 1
	EDT001 1		SPC001 1

The clones (identified by internal designations numbers, used to deduce the coding region for CG007 are set out in Table 2 below.

Table 2
ESTs in the CG007alt1 Contig

RTA00002115.c.11	RTA00003518F.b.03.1	RTA00003676F.a.06.4
RTA00002115.c.06	RTA00003518F.b.09.1	RTA00003681F.b.07.4
RTA00002115.a.04	RTA00003518F.b.01.1	RTA00003555F.a.03.1
RTA00002114.a.02	RTA00003518F.b.07.1	RTA00003546F.a.01.1
RTA00002114.a.01	RTA00003677F.a.01.3	RTA00003518F.a.06.1
RTA00002107.a.01	RTA00003677F.a.02	RTA00003677F.a.06.1
RTA00002107.a.02	RTA00003518F.b.08.1	RTA00003634F.h.16.1
RTA00002095.e.21	RTA00003518F.b.14.1	RTA00003681F.c.07.2
RTA00001999.b.08	RTA00003518F.b.06.1	RTA00003681F.b.09.4
RTA00001399.h.07	RTA00003518F.a.22.1	RTA00002812F.h.16.1
RTA00000791.e.05	RTA00003518F.a.24.1	RTA00003548F.a.07.1
RTA00001806.m.22	RTA00002858F.n.10.2	RTA00003683F.d.21.2
RTA00001482.n.05	RTA00003518F.b.04.1	RTA00003676F.a.10.2
RTA00001809.n.17	RTA00003677F.a.05.3	RTA00003553F.a.08.1
RTA00000672.i.23	RTA00003518F.a.09.1	RTA00003676F.a.07.4
RTA00002104.e.09	RTA00002397F.d.17.1	RTA00003683F.a.03.2
RTA00001565.g.12	RTA00003676F.a.03.4	RTA00003683F.a.21.2
RTA00002188.d.14	RTA00002095F.e.21.1	RTA00003683F.a.11.2
RTA00002397.d.17	RTA00003522F.b.05.1	RTA00003682R.a.17.2
RTA00001880.l.16	RTA00003681F.a.21.4	RTA00003519F.a.01.1
RTA00003689F.a.03.1	RTA00003634F.h.15.1	RTA00003683F.b.01.2
RTA00003522F.b.18.1	RTA00003677F.a.02.1	RTA00003683F.a.19.2
RTA00003518F.a.17.1	RTA00002659F.a.02.1	RTA00003553F.b.11.1
RTA00003518F.b.05.1	RTA00003676F.a.08.4	RTA00003683F.c.01.2
RTA00003518F.b.17.1	RTA00002659F.a.03.1	RTA00003683F.d.06.2

RTA00003683F.a.15.2	RTA00003683F.d.20.2	RTA00003020F.a.02.1
RTA00003683F.a.24.2	RTA00003683F.a.09.2	RTA00003682F.a.18.1
RTA00003683F.b.03.2	RTA00003683F.b.12.2	RTA00003538F.a.04.1
RTA00003681F.c.02.2	RTA00003683F.b.14.2	RTA00003683F.d.19.2
RTA00003559F.a.03.2	RTA00003683F.a.17.2	RTA00003518F.a.15.1
RTA00003683F.a.22.2	RTA00003683F.b.07.2	RTA00003539F.a.01.1
RTA00003683F.d.18.2	RTA00003683F.a.07.2	RTA00003723F.k.10.1
RTA00003683F.b.02.2	RTA00003683F.b.10.2	RTA00003518F.a.10.1
RTA00003683F.c.02.2	RTA00003683F.a.16.2	RTA00003543F.a.05.3
RTA00003676F.a.02.4	RTA00003676F.a.10.4	RTA00003523F.a.01.1
RTA00003683F.b.24.2	RTA00003009F.a.21.1	RTA00003683F.a.08.2
RTA00003683F.b.06.2	RTA00003553F.a.20.1	RTA00003529F.a.11.3
RTA00003683F.a.12.2	RTA00003683F.d.15.2	RTA00003523F.a.02.1
RTA00003683F.b.04.2	RTA00003020F.a.01.1	RTA00003683F.c.23.2
RTA00003683F.b.05.2	RTA00003679F.a.16.1	RTA00003681F.b.10.4
RTA00003683F.a.18.2	RTA00003683F.b.11.2	RTA00002971F.e.08.1
RTA00003683F.a.14.2	RTA00003683F.b.16.2	RTA00003529F.a.11.3
RTA00003683F.d.08.2	RTA00003679F.b.19.1	RTA00003523F.a.02.1
RTA00003683F.a.20.2	RTA00003679F.c.14.1	RTA00003683F.c.23.2
		RTA00003681F.b.10.4

Clone CG015

CG015 was deduced by combining sequences CG015alt1 (internal designation tenascinAlt1_Hy030899, SEQ ID NO: 45) and CG015alt2 (internal designation tenascinAlt2_Hy040799, SEQ ID NO: 47). CG015alt1 was originally identified from library clones as set out in Table 3 and CG015alt2 was deduced from library clones set out in Table 4, both below.

Table 3
Library Clones for CG015alt1 (Tissue Distribution)

ABT004 1	LPC001 1
AKD001 1	LUC001 1
ALV002 1	AOV001 1
FLS001 2	

Table 4
Library Clones for CG015alt2 (Tissue Distribution)

ABT004 1	FLV001 1
AKD001 1	FSK001 1
ALV002 1	LPC001 1
AOV001 1	LUC001 1
FLS001 2	

CG015alt2 was identified in PCR reactions using a fetal skin library (five reactions including three primer pairs), a fetal lung library (three reactions including three primer pairs), and an adult brain library (one reaction). ESTs that were found to be common to both CG015alt1 andCG015alt2 included RTA00000242.c.08, RTA00002115.b.09, and RTA00002188.a.03.

Clone CG144

CG144, was identified by sequence CG144 (internal designation AngPO6_Hy061499, SEQ ID NO: 7. Library clones giving rise to the CG144 sequence included ESTs AHR001 1, and AOV001 10. No introns were identified in the CG144 sequence.

The complete CG0014 contig was deduced from ESTs having the internal designation numbers as set out in Table 5 below.

Table 5
ESTs in CG144 Contig

RTA00001506.h.16	RTA00003804.c.08
RTA00003011.p.06	RTA00003804.d.07
RTA00003015.d.01	RTA00003804.d.08
RTA00003804.a.04	RTA00003015.d.01
RTA00003804.a.11	RTA00003015.c.24
RTA00003804.b.02	RTA00003009.c.13
RTA00003804.b.08	RTA00003011.p.06
RTA00003804.c.07	RTA00002993.f.07
	RTA00002958.h.08

Clone CG250

CG250, encoded by sequence bearing internal designation tenascin2_Hy061199, SEQ ID NO: 9. No introns were identified in the coding region, which was identified from a single placental library clone. The sequence of CG250 also displays homology to polynucleotides previously identified that encode tenascin polypeptides.

EXAMPLE 2

5' RACE Extension of Angiopoietin Genes

Reaction conditions

5' RACE reactions were performed using two nested gene-specfic primers (GSP) and vector primers (VP) in sequential PCR reactions on a panel of cDNA libraries. The cDNA libraries used for RACE were prepared from mRNA using a random-primed, 5' capture method to enrich for the 5' ends of genes (Carninci et al, Genomics, 37, 327-336, 1996) and cloned into pSPORT vector (BRL Life Technologies) previously digested with NotI and SalI. The human mRNAs (Invitrogen) included message from adult brain, adult thymus, fetal muscle, fetal skin, fetal heart, fetal brain, fetal spleen, fetal liver, and fetal lung. In addition, adaptor-ligated cDNA pools (Marathon cDNAs, Clontech) made from human fetal kidney, fetal brain, adult ovary mRNAs were used in the RACE experiments.

In the first reaction, GSP1 (T_m ~80°C) and VP1 (T_m ~72°C) were mixed in a 5:1 ratio. Touchdown PCR was carried out as follows: an initial incubation at 96°C for one minute, followed by five cycles of 96°C for 30 seconds and 72°C for four minutes; five cycle of 96°C for 30 seconds and 70°C for four minutes; and 15 cycles of 96°C for 30 seconds and 68°C for four minutes. The products of the first reaction were diluted 1:20 and used as template for the second reaction. Primers GSP2 and VP2 (both T_m ~60°C) were mixed in a 1:1 ratio and PCR was carried out as follows: an initial incubation at 96°C for one minute; and 30 cycles of 96°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1:30 minutes. Final RACE products were separated and identified using agarose gel electrophoresis. Selected fragments were subcloned into a TA cloning vector and the inserts were sequenced.

For clone CG006, RACE was carried out using primers designed based on the sequence of CG006alt2. For clone CG007, primers were designed based on CG007alt1. For clone CG015, primers were designed based on the sequences of CG0015alt1 and CG015atl2. For clone CG144, primers were designed base on the only identified sequence. RACE was performed using all primer pairs described below with all libraries described above. The reactions that successfully provided extension of the various contigs are described below.

Vectors utilized in various library amplifications

A. Vector Primers

pSPORT VP1: 5' AGGCACCCCAGGCTTTACACTTTA SEQ ID NO: 15

3'pSPORT VP2: 5' TTCCCGGGTCGACGATTTCGT SEQ ID NO: 16

3'Marathon cDNA VP1: 5' CCATCCTAATACGACTCACTATAGGGC SEQ ID NO: 17

3'Marathon cDNA VP2: 5' ACTCACTATAGGGCTCGAGCGGC SEQ ID NO: 18

B. CG006GSPs and cDNAs used to complete the 5' end of Cg006alt2;1.

1. In human fetal liver cDNA2:

GSP1 (CG006R5): 5' GTCTTTCCAGTCTTCCAACTCAATTCGTA SEQ ID NO: 19

GSP2 (CG006R6): 5' GTATATCTTCTCTAGGCCCAA SEQ ID NO: 20

2. In human fetal liver cDNA3.

GSP1 (CG006R11): 5' GATGTTGAATTAATGTCCATGGACTACCTGAT SEQ ID NO: 21

GSP2 (CG006R10): 5' GGCATACATGCCACTTGTATGTT SEQ ID NO: 22

3. In human adult liver cDNA4.

GSP1 (CG006R12): 5' GATTTTGAATTAAGTTAGTTAGTTGCTCTTCTAAA SEQ ID NO: 23

GSP2 (CG006R13): 5' GAGTTGAGTTCAAGTGACATA SEQ ID NO: 24

4. In human adult liver

GSP1 (CG006R15): 5' TCATTAATTTGGCCCTTCGTCTTATGGACAAA

SEQ ID NO: 25

GSP2 (CG006R16): 5' GTCCCAACTGAAGGAGGCCAT SEQ ID NO: 26

5. RACE was carried out using fetal liver and adult liver libraries. Five reactions were carried out using the fetal liver library using two different primer pairs, all corresponding to the 3′ end of the coding region. Two reactions were carried out with the adult liver library using one primer pair corresponding to the 5′ end of the coding region. In fetal liver, the amplified sequence indicated that the intron at position 396 was not spliced out, but in the adult liver library, the fully processed message was present.

In CG006alt2, introns were identified at positions 396, 927, and 1021. The exon between 927 and 1021 was found to be unique to CG006alt2, as it appeared to be spliced out of CG006alt3.

C. cDNACG007GSPs and cDNAs used to complete the 5' end of CG007:1.

1. In human fetal muscle cDNA2

GSP1 (CG007R1): 5' GCAGGCTATATGCCGTGTTCTCGCCACCA

SEQ ID NO: 27

GSP2 (CG007R2): 5' CCCGCAGTTGCACGGCCAGGC SEQ ID NO: 28

2. In human fetal muscle cDNA, human fetal brain cDNA, and human fetal skin cDNA3.

GSP1 (CG007R5): 5' TGCTGAATTCGCAGGTGCTGCTT SEQ ID NO: 29

GSP2 (CG007R6): 5' GCTGGGCCACCTTGTGGA SEQ ID NO: 30

3. In human fetal muscle cDNA, human fetal brain cDNA, human fetal skin cDNA, and human fetal kidney

GSP1 (CG007R7): 5' CTGCAGGAGTCCGTGCGCCAGGACATT

SEQ ID NO: 31

GSP2 (CG007R8): 5' ATCTCGTCCCAGGACGCAAA SEQ ID NO: 32

4. RACE was carried out using four libraries. In fetal brain, three reactions were performed using two primer pairs; in fetal kidney, one reaction was carried out; in fetal

skin, two reactions were carried out using two primer pairs; and in fetal muscle, three reactions were performed using three primer pairs.

D. cDNACG144GSPs and cDNAs used to complete the 5' end of Cg144:1.

1. In human ovary cDNA2

GSP1 (CG144R1): 5' CCATGTGACTGAACAGGTCTGTGAGGAAAA SEQ ID NO: 33

GSP2 (CG144R2): 5' GAACTCTATTCATGAGCTCGTTA SEQ ID NO: 34

2. In human ovary

GSP1 (CG144R3): 5' ACATGATTCCTCACAGTCTTCCTTACAAA SEQ ID NO: 35

GSP2 (CG144R4): 5' ACTACTGAAGAGTCCGTAGAA SEQ ID NO: 36

3. RACE was performed using an adult ovary library in a single reaction.

E. cDNACG015GSPs and cDNAs used to extend the 5' end of both CG015alt1 and CG015alt2:1.

1. In human fetal skin cDNA and human fetal lung cDNA2.

GSP1 (CG015R1): 5' GAAAGAGAGTCTCCAGCATCACCTACCAT SEQ ID NO: 37

GSP2 (CG015R3): 5' CCAGGGAGAAGCCATCATAGT SEQ ID NO: 38

2. In human fetal skin cDNA and human fetal liver cDNA3.

GSP1 (CG015alt1R5): 5' GGCTCTGGGGCTGGGTCCAGCATCCTA SEQ ID NO: 39

GSP2 (CG015alt1R6): 5' ACCCACAAGACGGACCGGAA SEQ ID NO: 40

3. In human fetal skin cDNA, human fetal lung cDNA, and human fetal liver cDNA4.

GSP1 (CG015alt2R5): 5' GGGTGACCTGCAGGCATGGGAGAAGCAT SEQ ID NO: 41

GSP2 (CG015alt2R6): 5' GGCTGGGTCCAGCATCCTA SEQ ID NO: 42

4. In human fetal skin cDNA

GSP1 (CG015alt1R5): 5' GGCTCTGGGGCTGGGTCCAGCATCCTA SEQ ID NO: 43

GSP2 (CG015alt1R7): 5' GTGGCGGCAGGACCTGCT SEQ ID NO: 44

5. RACE was carried out for CG015alt1 using a fetal skin library (four reactions using three primer pairs), a fetal lung library (two reactions including two primer pairs), and fetal liver (one reaction).

Sequences

RACE permitted extension of the 5' ends of clones CG006, CG007, CG015, and CG144. Based on the sequences of the underlying ESTs to define the gene and the sequences identified by RACE, the polynucleotide and amino acid sequences for CG006, CG007, CG015, and CG144 are set out as described above. Using the Signal P sequence analysis program, potential signal sequences were determined for proteins encoded by CG007 and CG144. In the CG007 protein, the signal sequence is predicted to be amino acid residues 1 through 25 as set out in SEQ ID NO: 6, and for the protein encoded by CG144, the signal is predicted to be amino acid residues 1 through 22 in SEQ ID NO: 8. The complete sequence for CG006alt2 was found to be identical to Ang5, a new angiopoietin entered into Genbank May 18, 1999, Accession Number AF152562_1. The sequence for CG006alt3, however, was determined to be distinct from all previously identified angiopoietins in that the CG006alt3clone lacked an exon found in the other sequences.

EXAMPLE 3

Identification of polymorphisms

Sequencing of a number of PCR products from various cDNA libraries is used to reveal potential polymorphisms, which are described with reference to the nucleotide sequence numbering of the SEQ ID NO: identified below. No polymorphisms were identified in CG006alt2, CG015alt1, CG144, or CG250.

Possible polymorphisms identified in CG007alt1 included G/A at position 296, T/G at position 777, and A/G at position 1216. For CG015alt2, numerous ESTs

were identified as suggesting polymorphisms for the gene sequence at A/C at position 521 and A/G at position 791.

EXAMPLE 4

Tissue Expression Study

PCR Analysis

Gene expression of the human angiopoietins is analyzed using a semi-quantitative PCR-based technique. A panel of cDNA libraries derived from human tissue (from Clontech and Invitrogen) is screened with angiopoietin specific primers to examine the mRNA expression of angiopoietin in human tissues and cell types. PCR assays (For example, 94 °C for 30 sec., 58 °C for 30 sec., 72 °C for 30 sec., for 30 cycles) are performed with 20 ng of cDNA derived from human tissues and cell lines and 10 picomoles of the angiopoietin gene-specific primers. The PCR product is identified through gel electrophoresis. Amplified products are separated on an agarose gel, transferred and chemically linked to a nylon filter. The filter is then hybridized with a radioactively labeled (^{33}Pα-dCTP) double-stranded probe generated from the full-length sequence using a Klenow polymerase, random prime method. The filters are washed (high stringency) and used to expose a phosphorimaging screen for several hours. Bands of the appropriate size indicate the presence of cDNA sequences in a specific library, and thus mRNA expression in the corresponding cell type or tissue.

Northern and Southern Analysis

Northern and Southern hybridizations were carried out in Church's buffer containing 7% SDS, 1% BSA, 1 mM EDTA, and 0.5 M $NaHPO_4$, pH 7.2. Hybridization was carried out at 65°C. Northerns were hybridized overnight, and RACE Southerns were hybridized from three hours to overnight. A final wash was carried out in 0.2X SSC/0.2% SDS at 65°C. Probes included purified PCR products amplified from cloned DNA labeled with ^{33}P-dATP (RACE Southerns) or ^{32}P-dATP (Northerns).

Northern results using a Clontech MTN blot indicated that CG006alt2 is expressed strongly in adult liver and weakly in adult kidney. There is a CG006alt2 major band at 1.6 kb and minor bands at 2.5 kb and 9.0 kb in both tissues. CG007 showed weak bands in fetal skin at 1.9 kb, fetal heart at 1.9 kb, and fetal kidney at 1.9 kb. CG015alt2

showed bands in fetal liver at 4.5 kb and 3.1 kb; fetal kidney at 4.5 kb and 3.1 kb; fetal brain at 4.5 kb; and fetal lung at 3.1 kb.

EXAMPLE 5

Chromosomal Localization Study

Chromosome mapping technologies allow investigators to link genes to specific regions of chromosomes. Chromosomal mapping is performed using the NIGMS human/rodent somatic cell hybrid mapping panel as described by Drwinga, H. L. et al., Genomics, 16, 311- 314, 1993 (human/rodent somatic cell hybrid mapping panel #2 purchased from the Coriell Institute for Medical Research, Camden, New Jersey). 60 ng of DNA from each sample in the panel is used as template, and 10 picomoles of the same angiopoietin gene-specific oligonucleotides are used as primers in a PCR assay (for example, 94°C for 30 sec., 58°C for 30 sec., 72°C for 30 sec., for 30 cycles). PCR products were analyzed by gel electrophoresis. The genomic PCR product is detected in a human/rodent somatic cell hybrid DNA containing a specific human chromosome.

By this technique, the CG006 angiopoietin gene has been mapped to chromosome 1, compared to the location of other known angiopoietins Angl at 8q22, Ang2 at 8p21, Ang4 at 20p13, and Ang5 at 1p31.1-p22.3.

EXAMPLE 6

Expression of Angiopoietin in E. coli

SEQ ID NO: 1, 3, 5, 7, 9, 11, 45, or 47 is expressed in E. coli by subcloning the entire coding region into a prokaryotic expression vector. The expression vector (pQE16) used is from the QIAexpression® prokaryotic protein expression system (QIAGEN). The features of this vector that make it useful for protein expression include: an efficient promoter (phage T5) to drive transcription; expression control provided by the lac operator system, which can be induced by addition of IPTG (isopropyl-β-D-thiogalactopyranoside), and an encoded His_6 tag. The latter is a stretch of 6 histidine amino acid residues which can bind very tightly to a nickel atom. The vector can be used to express a recombinant protein with a His_6 tag fused to its carboxyl terminus, allowing rapid and efficient purification using Ni-coupled affinity columns.

PCR is used to amplify the coding region which is then ligated into digested pQE16 vector. The ligation product is transformed by electroporation into electrocompetent E.coli cells (strain M15[pREP4] from QIAGEN), and the transformed cells are plated on ampicillin-containing plates. Colonies are screened for the correct insert in the proper orientation using a PCR reaction employing a gene-specific primer and a vector-specific primer. Positives are then sequenced to ensure correct orientation and sequence. To express angiopoietin, a colony containing a correct recombinant clone is inoculated into L-Broth containing 100 μg/ml of ampicillin, 25 μg/ml of kanamycin, and the culture was allowed to grow overnight at 37°C. The saturated culture is then diluted 20-fold in the same medium and allowed to grow to an optical density at 600 nm of 0.5. At this point, IPTG is added to a final concentration of 1 mM to induce protein expression. The culture is allowed to grow for 5 more hours, and then the cells are harvested by centrifugation at 3000 x g for 15 minutes.

The resultant pellet is lysed using a mild, nonionic detergent in 20 mM Tris HCl (pH 7.5) (B-PER™ Reagent from Pierce), or by sonication until the turbid cell suspension turned translucent. The lysate obtained is further purified using a nickel containing column (Ni-NTA spin column from QIAGEN) under non-denaturing conditions. Briefly, the lysate is brought up to 300 mM NaCl and 10 mM imidazole and centrifuged at 700 x g through the spin column to allow the His-tagged recombinant protein to bind to the nickel column. The column is then washed twice with Wash Buffer (50 mM NaH_2PO_4, pH 8.0; 300 mM NaCl; 20 mM imidazole) and is eluted with Elution Buffer (50 mM NaH_2PO_4, pH 8.0; 300 mM NaCl; 250 mM imidazole). All the above procedures are performed at 4°C. The presence of a purified protein of the predicted size is confirmed with SDS-PAGE.

E. coli Expression of CG007 and CG144

The following primers were utilized to amplify cDNA for clones CG007 and CG144. Reverse primers are shown as the reverse complement.

CG007 (Forward) #72 SEQ ID NO: 49
GAAAGATCTGGGACCCGTGCAGTCCAAGTC

CG007 (Reverse) #73 SEQ ID NO: 50
GCAGCAGAGGCAGCCTCCTAAGAAGAGCTCGAGAAGCTTGAC

CG144 (Forward) #90 SEQ ID NO: 51
GATTGATCAAGTACAAGGTAACTGTGTACATC

CG144 (Reverse) #91 SEQ ID NO: 52
GTACAATCCATATTTTAAATAATCTCGAGCCATGGATC

Templates for the PCR amplifications were cDNA clones containing full length open reading frames. Amplification was carried out using *pfu* polymerase. The retrieval numbers were as set out in Table 6.

Table 6

Clone	Retrieval	RTA Sequenced	Library
CG007	RTA3677.a.05	99999571a10	ADP001 - adipocyte
CG144	RTA3011.p.06	99999575a01	AOV001 - ovary

Individual amplification products were digested with the following enzymes: CG007, BglI and XhoI; and CG144 BclI and HindIII. In each digestion, the first restriction enzyme cleaved a site 5′ to the open reading frame and the second enzyme cleaved 3′ to the open reading frame. The digestion products were purified using low melting point agarose (FMC) and ligated into vector pRSETB previously digested with BamHI (which leaves an overhanging sequence compatible with BglI and BclI digestion products) and either XhoI or HindIII. The individual ligation mixtures were transformed into E. coli strain BL21(DE3)plysS, colonies were picked, and one of each was sequenced. The sequence of each selected clone was the same as previously deduced for the clones, with the exception of a silent change from CCG to CCA at position 1,303 of CG007.

In expression assays, transformed bacteria were grown at 37°C to OD 0.7 to 1.0 in 2XYT supplemented with 100 µl/ml carbenicillin and then induced with 1 mM IPTG. Following induction, cells were grown at 37°C or 25°C. Cells grown at 37°C were harvested at 3.5 hr post-induction and those grown at 25°C were harvested at 16 hr. None of the proteins were expressed at high levels and none were soluble.

EXAMPLE 7

Evaluation of Angiopoietin Activities *In Vitro and In Vivo*

Binding to the Tie-2 Receptor

A cell binding assay is carried out to demonstrate that angiopoietin polypeptides of the invention bind to the Tie-2 receptor. Briefly, cell binding of the recombinant protein with and without the presence of 100-fold greater amounts of non tagged angiopoietin ligand is analyzed by using fluorescent antibodies specific for a angiopoietin polypeptide (e.g. specific for an express tag within the recombinant polypeptide) on the fluorescent activated cell sorter (FACS). In each reaction, 10^6 cells NHDF (normal human dermal fibroblasts) are resuspended in 100 µl of FACS buffer (distilled PBS and 3% calf serum and 0.01% azide). Cell binding is done by adding 5 nM recombinant angiopoietin in 100 µl cell suspension and as a competition in one reaction, 500 nM of recombinant angiopoietin is also added. The cells are incubated on ice for 1 hr. The cells are pelleted, 200 µl of 0.2 mM BS3 (crosslinker) is added, and the cells are kept on ice for 30 min. Next, 10 µl 1 M Tris pH 7.5 is added and the cells are incubated for 15 minutes on ice. The cells are pelleted, washed 1 time in FACS buffer, resuspended in 100 µl volume of FACS buffer and 2 µl primary antibody (anti-express tag antibody 1 mg/ml) is added, and incubated on ice for 30 min. The cells are pelleted, washed with FACS buffer, and resuspended in FACS buffer (100 µl volume). The secondary antibody (phycoerythrin conjugated) 2 µl of anti-mouse Ig (1 mg/ml) is added and the cells are incubated for 30 minutes on ice. The cells are again pelleted, washed two times with FACS buffer, resuspended in 0.5 ml FACS buffer and analyzed on FACS. A shift in the fluorescence is expected to be observed in the cells treated with the recombinant tagged angiopoietin. This binding is shown to be specific if it is competed off with the non tagged angiopoietin protein.

The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention

are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims. All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a polynucleotide selected from the group consisting of:

(a) a polynucleotide having the nucleotide sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 14, 45, or 47;

(b) a polynucleotide having the angiopoietin protein coding nucleotide sequence of a polynucleotide of (a).

2. An isolated polynucleotide encoding a polypeptide with angiopoietin activity, comprising a polynucleotide that encodes the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48, or the mature protein sequence thereof;

3. An isolated polynucleotide encoding a polypeptide with angiopoietin activity that hybridizes under stringent conditions to the complement of a polynucleotide of any one of claims 1 or 2.

4. An isolated polynucleotide encoding a polypeptide with angiopoietin activity, said polynucleotide having greater than about 80% sequence identity with the polynucleotide of claim 1 or 2.

5. The polynucleotide of claim 1 or 2 which is a DNA.

6. An isolated polynucleotide which comprises a complement of the polynucleotide of claim 1.

7. An expression vector comprising the DNA of claim 5.

8. A host cell genetically engineered to contain the DNA of claim 5.

9. A host cell genetically engineered to contain the DNA of claim 5 in operative association with a regulatory sequence that controls expression of the DNA in the host cell.

10. An isolated polypeptide with angiopoietin activity comprising the angiopoietin protein sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 46, or 48, or the mature protein sequence thereof.

11. An isolated polypeptide with angiopoietin activity selected from the group consisting

a) a polypeptide having greater than about 80% sequence identity with the polypeptide of claim 10, and

b) a polypeptide encoded by the polynucleotide of claim 3.

12. A composition comprising the polypeptide of claim 10 or 11 and a carrier.

13. An antibody directed against the polypeptide of claim 10 or 11.

14. A method for detecting a polynucleotide of claim 3 in a sample, comprising the steps of:

a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex; and

b) detecting the complex, so that if a complex is detected, a polynucleotide of claim 1 is detected.

15. A method for detecting a polynucleotide of claim 3 in a sample, comprising the steps of:

a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of claim 1 under such conditions; and

b) amplifying the polynucleotides of claim 1 so that if a polynucleotide is amplified, a polynucleotide of claim 1 is detected.

16. The method of claim 15, wherein the polynucleotide is an RNA molecule that encodes a polypeptide of claim 11, and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.

17. A method for detecting a polypeptide of claim 11 in a sample, comprising:

a) contacting the sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex; and

b) detecting the complex, so that if a complex is detected, a polypeptide of claim 11 is detected.

18. A method for identifying a compound that binds to a polypeptide of claim 11, comprising:

a) contacting a compound with a polypeptide of claim 11 for a time sufficient to form a polypeptide/compound complex; and

b) detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polypeptide of claim 11 is identified.

19. A method for identifying a compound that binds to a polypeptide of claim 11, comprising:

a) contacting a compound with a polypeptide of claim 11, in a cell, for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

b) detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds to a polypeptide of claim 11 is identified.

20. A method of producing the polypeptide of claim 11, comprising,

a) culturing the host cell of claim 9 for a period of time sufficient to express the polypeptide; and

b) isolating the polypeptide from the cell or culture media in which the cell is grown.

21. A kit comprising the polypeptide of claim 11.

SEQUENCE LISTING

<110> Ballinger, Dennis

<120> Novel Angiopoietins Materials and Methods

<130> 28110/36597

<140>
<141>

<150> 09/354,881
<151> 1999-07-16

<160> 52

<170> PatentIn Ver. 2.0

<210> 1
<211> 1610
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (64)..(1443)

<220>
<223> CG006-alt2

<400> 1
gtctaggtct gcttccagaa gaaaacagtt ccacgttgct tgaaattgaa aatcaagata 60

aaa atg ttc aca att aag ctc ctt ctt ttt att gtt cct cta gtt att 108
Met Phe Thr Ile Lys Leu Leu Leu Phe Ile Val Pro Leu Val Ile
1 5 10 15

tcc tcc aga att gat caa gac aat tca tca ttt gat tct cta tct cca 156
Ser Ser Arg Ile Asp Gln Asp Asn Ser Ser Phe Asp Ser Leu Ser Pro
20 25 30

gag cca aaa tca aga ttt gct atg tta gac gat gta aaa att tta gcc 204
Glu Pro Lys Ser Arg Phe Ala Met Leu Asp Asp Val Lys Ile Leu Ala
35 40 45

aat ggc ctc ctt cag ttg gga cat ggt ctt aaa gac ttt gtc cat aag 252
Asn Gly Leu Leu Gln Leu Gly His Gly Leu Lys Asp Phe Val His Lys
50 55 60

acg aag ggc caa att aat gac ata ttt caa aaa ctc aac ata ttt gat 300
Thr Lys Gly Gln Ile Asn Asp Ile Phe Gln Lys Leu Asn Ile Phe Asp
65 70 75

cag tct ttt tat gat cta tcg ctg caa acc agt gaa atc aaa gaa gaa 348
Gln Ser Phe Tyr Asp Leu Ser Leu Gln Thr Ser Glu Ile Lys Glu Glu
80 85 90 95

gaa aag gaa ctg aga aga act aca tat aaa cta caa gtc aaa aat gaa 396
Glu Lys Glu Leu Arg Arg Thr Thr Tyr Lys Leu Gln Val Lys Asn Glu
100 105 110

```
gag gta aag aat atg tca ctt gaa ctc aac tca aaa ctt gaa agc ctc   444
Glu Val Lys Asn Met Ser Leu Glu Leu Asn Ser Lys Leu Glu Ser Leu
            115                 120                 125

cta gaa gaa aaa att cta ctt caa caa aaa gtg aaa tat tta gaa gag   492
Leu Glu Glu Lys Ile Leu Leu Gln Gln Lys Val Lys Tyr Leu Glu Glu
        130                 135                 140

caa cta act aac tta att caa aat caa cct gga act cca gaa cac cca   540
Gln Leu Thr Asn Leu Ile Gln Asn Gln Pro Gly Thr Pro Glu His Pro
    145                 150                 155

gaa gta act tca ctt aaa act ttt gta gaa aaa caa gat aat agc atc   588
Glu Val Thr Ser Leu Lys Thr Phe Val Glu Lys Gln Asp Asn Ser Ile
160                 165                 170                 175

aaa gac ctt ctc cag acc gtg gaa gac caa tat aaa caa tta aac caa   636
Lys Asp Leu Leu Gln Thr Val Glu Asp Gln Tyr Lys Gln Leu Asn Gln
                180                 185                 190

cag cat agt caa ata aaa gaa ata gaa aat cag ctc aga agg act agt   684
Gln His Ser Gln Ile Lys Glu Ile Glu Asn Gln Leu Arg Arg Thr Ser
            195                 200                 205

att caa gaa ccc aca gaa att tct cta tct tcc aag cca aga gca cca   732
Ile Gln Glu Pro Thr Glu Ile Ser Leu Ser Ser Lys Pro Arg Ala Pro
        210                 215                 220

aga act act ccc ttt ctt cag ttg aat gaa ata aga aat gta aaa cat   780
Arg Thr Thr Pro Phe Leu Gln Leu Asn Glu Ile Arg Asn Val Lys His
    225                 230                 235

gat ggc att cct gct gaa tgt acc acc att tat aac aga ggt gaa cat   828
Asp Gly Ile Pro Ala Glu Cys Thr Thr Ile Tyr Asn Arg Gly Glu His
240                 245                 250                 255

aca agt ggc atg tat gcc atc aga ccc agc aac tct caa gtt ttt cat   876
Thr Ser Gly Met Tyr Ala Ile Arg Pro Ser Asn Ser Gln Val Phe His
                260                 265                 270

gtc tac tgt gat gtt ata tca ggt agt cca tgg aca tta att caa cat   924
Val Tyr Cys Asp Val Ile Ser Gly Ser Pro Trp Thr Leu Ile Gln His
            275                 280                 285

cga ata gat gga tca caa aac ttc aat gaa acg tgg gag aac tac aaa   972
Arg Ile Asp Gly Ser Gln Asn Phe Asn Glu Thr Trp Glu Asn Tyr Lys
        290                 295                 300

tat ggt ttt ggg agg ctt gat gga gaa ttt tgg ttg ggc cta gag aag  1020
Tyr Gly Phe Gly Arg Leu Asp Gly Glu Phe Trp Leu Gly Leu Glu Lys
    305                 310                 315

ata tac tcc ata gtg aag caa tct aat tat gtt tta cga att gag ttg  1068
Ile Tyr Ser Ile Val Lys Gln Ser Asn Tyr Val Leu Arg Ile Glu Leu
320                 325                 330                 335

gaa gac tgg aaa gac aac aaa cat tat att gaa tat tct ttt tac ttg  1116
Glu Asp Trp Lys Asp Asn Lys His Tyr Ile Glu Tyr Ser Phe Tyr Leu
                340                 345                 350
```

```
gga aat cac gaa acc aac tat acg cta cat cta gtt gcg att act ggc   1164
Gly Asn His Glu Thr Asn Tyr Thr Leu His Leu Val Ala Ile Thr Gly
            355                 360                 365

aat gtc ccc aat gca atc ccg gaa aac aaa gat ttg gtg ttt tct act   1212
Asn Val Pro Asn Ala Ile Pro Glu Asn Lys Asp Leu Val Phe Ser Thr
        370                 375                 380

tgg gat cac aaa gca aaa gga cac ttc aac tgt cca gag ggt tat tca   1260
Trp Asp His Lys Ala Lys Gly His Phe Asn Cys Pro Glu Gly Tyr Ser
    385                 390                 395

gga ggc tgg tgg tgg cat gat gag tgt gga gaa aac aac cta aat ggt   1308
Gly Gly Trp Trp Trp His Asp Glu Cys Gly Glu Asn Asn Leu Asn Gly
400                 405                 410                 415

aaa tat aac aaa cca aga gca aaa tct aag cca gag agg aga aga gga   1356
Lys Tyr Asn Lys Pro Arg Ala Lys Ser Lys Pro Glu Arg Arg Arg Gly
                420                 425                 430

tta tct tgg aag tct caa aat gga agg tta tac tct ata aaa tca acc   1404
Leu Ser Trp Lys Ser Gln Asn Gly Arg Leu Tyr Ser Ile Lys Ser Thr
            435                 440                 445

aaa atg ttg atc cat cca aca gat tca gaa agc ttt gaa tgaactgagg    1453
Lys Met Leu Ile His Pro Thr Asp Ser Glu Ser Phe Glu
        450                 455                 460

caaatttaaa aggcaataat ttaaacatta acctcattcc aagttaatgt ggtctaataa 1513

tctggtatta aatccttaag agaaagcttg agaaatagat ttttttttatc ttaaagtcac 1573

tgtctattta agattaaaca tacaatcaca taaccttt                          1610
```

<210> 2
<211> 460
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Phe Thr Ile Lys Leu Leu Leu Phe Ile Val Pro Leu Val Ile Ser
  1               5                  10                  15

Ser Arg Ile Asp Gln Asp Asn Ser Ser Phe Asp Ser Leu Ser Pro Glu
             20                  25                  30

Pro Lys Ser Arg Phe Ala Met Leu Asp Asp Val Lys Ile Leu Ala Asn
         35                  40                  45

Gly Leu Leu Gln Leu Gly His Gly Leu Lys Asp Phe Val His Lys Thr
     50                  55                  60

Lys Gly Gln Ile Asn Asp Ile Phe Gln Lys Leu Asn Ile Phe Asp Gln
 65                  70                  75                  80

Ser Phe Tyr Asp Leu Ser Leu Gln Thr Ser Glu Ile Lys Glu Glu Glu
                 85                  90                  95

Lys Glu Leu Arg Arg Thr Thr Tyr Lys Leu Gln Val Lys Asn Glu Glu
            100                 105                 110
```

```
Val Lys Asn Met Ser Leu Glu Leu Asn Ser Lys Leu Glu Ser Leu Leu
        115                 120                 125

Glu Glu Lys Ile Leu Leu Gln Gln Lys Val Lys Tyr Leu Glu Glu Gln
    130                 135                 140

Leu Thr Asn Leu Ile Gln Asn Gln Pro Gly Thr Pro Glu His Pro Glu
145                 150                 155                 160

Val Thr Ser Leu Lys Thr Phe Val Glu Lys Gln Asp Asn Ser Ile Lys
                165                 170                 175

Asp Leu Leu Gln Thr Val Glu Asp Gln Tyr Lys Gln Leu Asn Gln Gln
            180                 185                 190

His Ser Gln Ile Lys Glu Ile Glu Asn Gln Leu Arg Arg Thr Ser Ile
        195                 200                 205

Gln Glu Pro Thr Glu Ile Ser Leu Ser Ser Lys Pro Arg Ala Pro Arg
    210                 215                 220

Thr Thr Pro Phe Leu Gln Leu Asn Glu Ile Arg Asn Val Lys His Asp
225                 230                 235                 240

Gly Ile Pro Ala Glu Cys Thr Thr Ile Tyr Asn Arg Gly Glu His Thr
                245                 250                 255

Ser Gly Met Tyr Ala Ile Arg Pro Ser Asn Ser Gln Val Phe His Val
            260                 265                 270

Tyr Cys Asp Val Ile Ser Gly Ser Pro Trp Thr Leu Ile Gln His Arg
        275                 280                 285

Ile Asp Gly Ser Gln Asn Phe Asn Glu Thr Trp Glu Asn Tyr Lys Tyr
    290                 295                 300

Gly Phe Gly Arg Leu Asp Gly Glu Phe Trp Leu Gly Leu Glu Lys Ile
305                 310                 315                 320

Tyr Ser Ile Val Lys Gln Ser Asn Tyr Val Leu Arg Ile Glu Leu Glu
                325                 330                 335

Asp Trp Lys Asp Asn Lys His Tyr Ile Glu Tyr Ser Phe Tyr Leu Gly
            340                 345                 350

Asn His Glu Thr Asn Tyr Thr Leu His Leu Val Ala Ile Thr Gly Asn
        355                 360                 365

Val Pro Asn Ala Ile Pro Glu Asn Lys Asp Leu Val Phe Ser Thr Trp
    370                 375                 380

Asp His Lys Ala Lys Gly His Phe Asn Cys Pro Glu Gly Tyr Ser Gly
385                 390                 395                 400

Gly Trp Trp Trp His Asp Glu Cys Gly Glu Asn Asn Leu Asn Gly Lys
                405                 410                 415

Tyr Asn Lys Pro Arg Ala Lys Ser Lys Pro Glu Arg Arg Arg Gly Leu
            420                 425                 430

Ser Trp Lys Ser Gln Asn Gly Arg Leu Tyr Ser Ile Lys Ser Thr Lys
        435                 440                 445
```

```
Met Leu Ile His Pro Thr Asp Ser Glu Ser Phe Glu
    450                 455                 460


<210> 3
<211> 699
<212> DNA
<213> Homo sapiens

<220>
<223> CG006-alt3

<220>
<221> CDS
<222> (2)..(520)

<400> 3
t aca agt ggc atg tat gcc atc aga ccc agc aac tct caa gtt ttt cat 49
  Thr Ser Gly Met Tyr Ala Ile Arg Pro Ser Asn Ser Gln Val Phe His
    1               5                  10                  15

gtc tac tgg gat gtt ata tca gga gaa ttt tgg ttg ggc cta gag aag   97
Val Tyr Trp Asp Val Ile Ser Gly Glu Phe Trp Leu Gly Leu Glu Lys
            20                  25                  30

ata tac tcc ata gtg aag caa tct aat tat gtt tta cga att gag ttg   145
Ile Tyr Ser Ile Val Lys Gln Ser Asn Tyr Val Leu Arg Ile Glu Leu
        35                  40                  45

gaa gac tgg aaa gac aac aaa cat tat att gaa tat tct ttt tac ttg   193
Glu Asp Trp Lys Asp Asn Lys His Tyr Ile Glu Tyr Ser Phe Tyr Leu
    50                  55                  60

gga aat cac gaa acc aac tat acg cta cat cta gtt gcg att act ggc   241
Gly Asn His Glu Thr Asn Tyr Thr Leu His Leu Val Ala Ile Thr Gly
65                  70                  75                  80

aat gtc ccc aat gca atc ccg gaa aac aaa gat ttg gtg ttt tct act   289
Asn Val Pro Asn Ala Ile Pro Glu Asn Lys Asp Leu Val Phe Ser Thr
                85                  90                  95

tgg gat cac aaa gca aaa gga cac ttc aac tgt cca gag ggt tat tca   337
Trp Asp His Lys Ala Lys Gly His Phe Asn Cys Pro Glu Gly Tyr Ser
            100                 105                 110

gga ggc tgg tgg tgg cat gat gag tgt gga gaa aac aac cta aat ggt   385
Gly Gly Trp Trp Trp His Asp Glu Cys Gly Glu Asn Asn Leu Asn Gly
        115                 120                 125

aaa tat aac aaa cca aga gca aaa tct aag cca gag agg aga aga gga   433
Lys Tyr Asn Lys Pro Arg Ala Lys Ser Lys Pro Glu Arg Arg Arg Gly
    130                 135                 140

tta tct tgg aag tct caa aat gga agg tta tac tct ata aaa tca acc   481
Leu Ser Trp Lys Ser Gln Asn Gly Arg Leu Tyr Ser Ile Lys Ser Thr
145                 150                 155                 160

aaa atg ttg atc cat cca aca gat tca gaa agc ttt gaa tgaactgagg    530
Lys Met Leu Ile His Pro Thr Asp Ser Glu Ser Phe Glu
                165                 170
```

caaatttaaa aggcaataat ttaaacatta acctcattcc aagttaatgt ggtctaataa 590

tctggtatta aatccttaag agaaagcttg agaaatagat ttttttatc ttaaagtcac 650

tgtctattta agattaaaca tacaatcaca taaccttaaa aaaaaaaaa 699

<210> 4
<211> 173
<212> PRT
<213> Homo sapiens

<400> 4

Thr	Ser	Gly	Met	Tyr	Ala	Ile	Arg	Pro	Ser	Asn	Ser	Gln	Val	Phe	His
1				5				10					15		
Val	Tyr	Trp	Asp	Val	Ile	Ser	Gly	Glu	Phe	Trp	Leu	Gly	Leu	Glu	Lys
			20					25					30		
Ile	Tyr	Ser	Ile	Val	Lys	Gln	Ser	Asn	Tyr	Val	Leu	Arg	Ile	Glu	Leu
		35					40					45			
Glu	Asp	Trp	Lys	Asp	Asn	Lys	His	Tyr	Ile	Glu	Tyr	Ser	Phe	Tyr	Leu
	50					55					60				
Gly	Asn	His	Glu	Thr	Asn	Tyr	Thr	Leu	His	Leu	Val	Ala	Ile	Thr	Gly
65				70					75					80	
Asn	Val	Pro	Asn	Ala	Ile	Pro	Glu	Asn	Lys	Asp	Leu	Val	Phe	Ser	Thr
			85					90					95		
Trp	Asp	His	Lys	Ala	Lys	Gly	His	Phe	Asn	Cys	Pro	Glu	Gly	Tyr	Ser
		100					105					110			
Gly	Gly	Trp	Trp	Trp	His	Asp	Glu	Cys	Gly	Glu	Asn	Asn	Leu	Asn	Gly
	115					120					125				
Lys	Tyr	Asn	Lys	Pro	Arg	Ala	Lys	Ser	Lys	Pro	Glu	Arg	Arg	Arg	Gly
130					135					140					
Leu	Ser	Trp	Lys	Ser	Gln	Asn	Gly	Arg	Leu	Tyr	Ser	Ile	Lys	Ser	Thr
145				150					155					160	
Lys	Met	Leu	Ile	His	Pro	Thr	Asp	Ser	Glu	Ser	Phe	Glu			
				165					170						

<210> 5
<211> 1862
<212> DNA
<213> Homo sapiens

<220>
<223> CG007

<220>
<221> CDS
<222> (137)..(1354)

<400> 5
tgtgatccga ttctttccag cggcttctgc aaccaagcgg gtcttacccc cggtcctccg 60

```
cgtctccagt cctcgcacct ggaaccccaa cgtccccgag agtccccgaa tccccgctcc 120

caggctacct aagagg atg agc ggt gct ccg acg gcc ggg gca gcc ctg atg 172
                Met Ser Gly Ala Pro Thr Ala Gly Ala Ala Leu Met
                  1               5                  10

ctc tgc gcc gcc acc gcc gtg cta ctg agc gct cag ggc gga ccc gtg   220
Leu Cys Ala Ala Thr Ala Val Leu Leu Ser Ala Gln Gly Gly Pro Val
         15                  20                  25

cag tcc aag tcg ccg cgc ttt gcg tcc tgg gac gag atg aat gtc ctg   268
Gln Ser Lys Ser Pro Arg Phe Ala Ser Trp Asp Glu Met Asn Val Leu
     30                  35                  40

gcg cac gga ctc ctg cag ctc ggc cag ggg ctg cgc gaa cac gcg gag   316
Ala His Gly Leu Leu Gln Leu Gly Gln Gly Leu Arg Glu His Ala Glu
 45                  50                  55                  60

cgc acc cgc agt cag ctg agc gcg ctg gag cgg cgc ctg agc gcg tgc   364
Arg Thr Arg Ser Gln Leu Ser Ala Leu Glu Arg Arg Leu Ser Ala Cys
                 65                  70                  75

ggg tcc gcc tgt cag gga acc gag ggg tcc acc gac ctc ccg tta gcc   412
Gly Ser Ala Cys Gln Gly Thr Glu Gly Ser Thr Asp Leu Pro Leu Ala
             80                  85                  90

cct gag agc cgg gtg gac cct gag gtc ctt cac agc ctg cag aca caa   460
Pro Glu Ser Arg Val Asp Pro Glu Val Leu His Ser Leu Gln Thr Gln
         95                  100                 105

ctc aag gct cag aac agc agg atc cag caa ctc ttc cac aag gtg gcc   508
Leu Lys Ala Gln Asn Ser Arg Ile Gln Gln Leu Phe His Lys Val Ala
    110                 115                 120

cag cag cag cgg cac ctg gag aag cag cac ctg cga att cag cat ctg   556
Gln Gln Gln Arg His Leu Glu Lys Gln His Leu Arg Ile Gln His Leu
125                 130                 135                 140

caa agc cag ttt ggc ctc ctg gac cac aag cac cta gac cat gag gtg   604
Gln Ser Gln Phe Gly Leu Leu Asp His Lys His Leu Asp His Glu Val
                145                 150                 155

gcc aag cct gcc cga aga aag agg ctg ccc gag atg gcc cag cca gtt   652
Ala Lys Pro Ala Arg Arg Lys Arg Leu Pro Glu Met Ala Gln Pro Val
            160                 165                 170

gac ccg gct cac aat gtc agc cgc ctg cac cgg ctg ccc agg gat tgc   700
Asp Pro Ala His Asn Val Ser Arg Leu His Arg Leu Pro Arg Asp Cys
        175                 180                 185

cag gag ctg ttc cag gtt ggg gag agg cag agt gga cta ttt gaa atc   748
Gln Glu Leu Phe Gln Val Gly Glu Arg Gln Ser Gly Leu Phe Glu Ile
    190                 195                 200

cag cct cag ggg tct ccg cca ttt ttg gtg aac tgc aag atg acc tca   796
Gln Pro Gln Gly Ser Pro Pro Phe Leu Val Asn Cys Lys Met Thr Ser
205                 210                 215                 220

gat gga ggc tgg aca gta att cag agg cgc cac gat ggc tca gtg gac   844
Asp Gly Gly Trp Thr Val Ile Gln Arg Arg His Asp Gly Ser Val Asp
                225                 230                 235
```

```
ttc aac cgg ccc tgg gaa gcc tac aag gcg ggg ttt ggg gat ccc cac   892
Phe Asn Arg Pro Trp Glu Ala Tyr Lys Ala Gly Phe Gly Asp Pro His
            240                 245                 250

ggc gag ttc tgg ctg ggt ctg gag aag gtg cat agc atc acg ggg gac   940
Gly Glu Phe Trp Leu Gly Leu Glu Lys Val His Ser Ile Thr Gly Asp
        255                 260                 265

cgc aac agc cgc ctg gcc gtg cag ctg cgg gac tgg gat ggc aac gcc   988
Arg Asn Ser Arg Leu Ala Val Gln Leu Arg Asp Trp Asp Gly Asn Ala
    270                 275                 280

gag ttg ctg cag ttc tcc gtg cac ctg ggt ggc gag gac acg gcc tat   1036
Glu Leu Leu Gln Phe Ser Val His Leu Gly Gly Glu Asp Thr Ala Tyr
285                 290                 295                 300

agc ctg cag ctc act gca ccc gtg gcc ggc cag ctg ggc gcc acc acc   1084
Ser Leu Gln Leu Thr Ala Pro Val Ala Gly Gln Leu Gly Ala Thr Thr
                305                 310                 315

gtc cca ccc agc ggc ctc tcc gta ccc ttc tcc act tgg gac cag gat   1132
Val Pro Pro Ser Gly Leu Ser Val Pro Phe Ser Thr Trp Asp Gln Asp
            320                 325                 330

cac gac ctc cgc agg gac aag aac tgc gcc aag agc ctc tct gga ggc   1180
His Asp Leu Arg Arg Asp Lys Asn Cys Ala Lys Ser Leu Ser Gly Gly
        335                 340                 345

tgg tgg ttt ggc acc tgc agc cat tcc aac ctc aac ggc cag tac ttc   1228
Trp Trp Phe Gly Thr Cys Ser His Ser Asn Leu Asn Gly Gln Tyr Phe
    350                 355                 360

cgc tcc atc cca cag cag cgg cag aag ctt aag aag gga atc ttc tgg   1276
Arg Ser Ile Pro Gln Gln Arg Gln Lys Leu Lys Lys Gly Ile Phe Trp
365                 370                 375                 380

aag acc tgg cgg ggc cgc tac tac ccg ctg cag gcc acc acc atg ttg   1324
Lys Thr Trp Arg Gly Arg Tyr Tyr Pro Leu Gln Ala Thr Thr Met Leu
                385                 390                 395

atc cag ccc atg gca gca gag gca gcc tcc tagcgtcctg gctgggcctg     1374
Ile Gln Pro Met Ala Ala Glu Ala Ala Ser
            400                 405

gtcccaggcc cacgaaagac ggtgactctt ggctctgccc gaggatgtgg ccgttccctg 1434

cctgggcagg ggctccaagg aggggccatc tggaaacttg tggacagaga agaagaccac 1494

gactggagaa gccccctttc tgagtgcagg ggggctgcat gcgttgcctc ctgagatcga 1554

ggctgcagga tatgctcaga ctctagaggc gtggaccaag gggcatggag cttcactcct 1614

tgctggccag ggagttgggg actcagaggg accacttggg gccagccaga ctggcctcaa 1674

tggcggactc agtcacattg actgacgggg accagggctt gtgtgggtcg agagcgccct 1734

catggtgctg gtgctgttgt gtgtaggtcc cctggggaca caagcaggcg ccaatggtat 1794

ctgggcggag ctcacagagt tcttggaata aaagcaacct cagaacaaaa aaaaaaaaaa 1854

aaaaaaaa                                                          1862
```

<210> 6
<211> 406
<212> PRT
<213> Homo sapiens

<400> 6

```
Met Ser Gly Ala Pro Thr Ala Gly Ala Ala Leu Met Leu Cys Ala Ala
  1               5                  10                  15

Thr Ala Val Leu Leu Ser Ala Gln Gly Gly Pro Val Gln Ser Lys Ser
             20                  25                  30

Pro Arg Phe Ala Ser Trp Asp Glu Met Asn Val Leu Ala His Gly Leu
         35                  40                  45

Leu Gln Leu Gly Gln Gly Leu Arg Glu His Ala Glu Arg Thr Arg Ser
     50                  55                  60

Gln Leu Ser Ala Leu Glu Arg Arg Leu Ser Ala Cys Gly Ser Ala Cys
 65                  70                  75                  80

Gln Gly Thr Glu Gly Ser Thr Asp Leu Pro Leu Ala Pro Glu Ser Arg
                 85                  90                  95

Val Asp Pro Glu Val Leu His Ser Leu Gln Thr Gln Leu Lys Ala Gln
            100                 105                 110

Asn Ser Arg Ile Gln Gln Leu Phe His Lys Val Ala Gln Gln Gln Arg
        115                 120                 125

His Leu Glu Lys Gln His Leu Arg Ile Gln His Leu Gln Ser Gln Phe
    130                 135                 140

Gly Leu Leu Asp His Lys His Leu Asp His Glu Val Ala Lys Pro Ala
145                 150                 155                 160

Arg Arg Lys Arg Leu Pro Glu Met Ala Gln Pro Val Asp Pro Ala His
                165                 170                 175

Asn Val Ser Arg Leu His Arg Leu Pro Arg Asp Cys Gln Glu Leu Phe
            180                 185                 190

Gln Val Gly Glu Arg Gln Ser Gly Leu Phe Glu Ile Gln Pro Gln Gly
        195                 200                 205

Ser Pro Pro Phe Leu Val Asn Cys Lys Met Thr Ser Asp Gly Gly Trp
    210                 215                 220

Thr Val Ile Gln Arg Arg His Asp Gly Ser Val Asp Phe Asn Arg Pro
225                 230                 235                 240

Trp Glu Ala Tyr Lys Ala Gly Phe Gly Asp Pro His Gly Glu Phe Trp
                245                 250                 255

Leu Gly Leu Glu Lys Val His Ser Ile Thr Gly Asp Arg Asn Ser Arg
            260                 265                 270

Leu Ala Val Gln Leu Arg Asp Trp Asp Gly Asn Ala Glu Leu Leu Gln
        275                 280                 285

Phe Ser Val His Leu Gly Gly Glu Asp Thr Ala Tyr Ser Leu Gln Leu
    290                 295                 300
```

```
Thr Ala Pro Val Ala Gly Gln Leu Gly Ala Thr Thr Val Pro Pro Ser
305                 310                 315                 320

Gly Leu Ser Val Pro Phe Ser Thr Trp Asp Gln Asp His Asp Leu Arg
                325                 330                 335

Arg Asp Lys Asn Cys Ala Lys Ser Leu Ser Gly Gly Trp Trp Phe Gly
            340                 345                 350

Thr Cys Ser His Ser Asn Leu Asn Gly Gln Tyr Phe Arg Ser Ile Pro
        355                 360                 365

Gln Gln Arg Gln Lys Leu Lys Lys Gly Ile Phe Trp Lys Thr Trp Arg
    370                 375                 380

Gly Arg Tyr Tyr Pro Leu Gln Ala Thr Thr Met Leu Ile Gln Pro Met
385                 390                 395                 400

Ala Ala Glu Ala Ala Ser
                405
```

<210> 7
<211> 1824
<212> DNA
<213> Homo sapiens

<220>
<223> CG144

<220>
<221> CDS
<222> (40)..(1203)

<220>
<223> Translation may initiate at the ATG codon at
nucleotides 40-42 or the ATG at nucleotides 43-45

<400> 7

```
tgatatttga agaagtgttt tcatctatcc aagaaaaat atg atg tct cca tcc    54
                                          Met Met Ser Pro Ser
                                            1               5

caa gcc tca ctc tta ttc tta aat gta tgt att ttt att tgt gga gaa   102
Gln Ala Ser Leu Leu Phe Leu Asn Val Cys Ile Phe Ile Cys Gly Glu
                10                  15                  20

gct gta caa ggt aac tgt gta cat cat tct acg gac tct tca gta gtt   150
Ala Val Gln Gly Asn Cys Val His His Ser Thr Asp Ser Ser Val Val
            25                  30                  35

aac att gta gaa gat gga tct aat gca aaa gat gaa agt aaa agt aat   198
Asn Ile Val Glu Asp Gly Ser Asn Ala Lys Asp Glu Ser Lys Ser Asn
        40                  45                  50

gat act gtt tgt aag gaa gac tgt gag gaa tca tgt gat gtt aaa act   246
Asp Thr Val Cys Lys Glu Asp Cys Glu Glu Ser Cys Asp Val Lys Thr
    55                  60                  65

aaa att aca cga gaa gaa aaa cat ttc atg tgt aga aat ttg caa aat   294
Lys Ile Thr Arg Glu Glu Lys His Phe Met Cys Arg Asn Leu Gln Asn
70                  75                  80                  85
```

```
tct att gtt tcc tac aca aga agt acc aaa aaa cta cta agg aat atg   342
Ser Ile Val Ser Tyr Thr Arg Ser Thr Lys Lys Leu Leu Arg Asn Met
                    90                  95                  100

atg gat gag caa caa gct tcc ttg gat tat tta tct aat cag gtt aac   390
Met Asp Glu Gln Gln Ala Ser Leu Asp Tyr Leu Ser Asn Gln Val Asn
            105                 110                 115

gag ctc atg aat aga gtt ctc ctt ttg act aca gaa gtt ttt aga aaa   438
Glu Leu Met Asn Arg Val Leu Leu Leu Thr Thr Glu Val Phe Arg Lys
        120                 125                 130

cag ctg gat cct ttt cct cac aga cct gtt cag tca cat ggt tta gat   486
Gln Leu Asp Pro Phe Pro His Arg Pro Val Gln Ser His Gly Leu Asp
    135                 140                 145

tgc act gat att aag gat acc att ggc tct gtc acc aaa aca ccg agt   534
Cys Thr Asp Ile Lys Asp Thr Ile Gly Ser Val Thr Lys Thr Pro Ser
150                 155                 160                 165

ggt tta tac ata att cac cca gaa gga tct agc tac cca ttt gag gta   582
Gly Leu Tyr Ile Ile His Pro Glu Gly Ser Ser Tyr Pro Phe Glu Val
                170                 175                 180

atg tgt gac atg gat tac aga gga ggt gga tgg act gtg ata cag aaa   630
Met Cys Asp Met Asp Tyr Arg Gly Gly Gly Trp Thr Val Ile Gln Lys
            185                 190                 195

aga att gat ggg ata att gat ttc cag agg ttg tgg tgt gat tat ctg   678
Arg Ile Asp Gly Ile Ile Asp Phe Gln Arg Leu Trp Cys Asp Tyr Leu
        200                 205                 210

gat gga ttt gga gat ctt cta gga gaa ttt tgg cta gga ctg aaa aag   726
Asp Gly Phe Gly Asp Leu Leu Gly Glu Phe Trp Leu Gly Leu Lys Lys
    215                 220                 225

att ttt tat ata gta aat cag aaa aat acc agt ttt atg ctg tat gtg   774
Ile Phe Tyr Ile Val Asn Gln Lys Asn Thr Ser Phe Met Leu Tyr Val
230                 235                 240                 245

gct ttg gaa tct gaa gat gac act ctt gct tat gca tca tat gat aat   822
Ala Leu Glu Ser Glu Asp Asp Thr Leu Ala Tyr Ala Ser Tyr Asp Asn
                250                 255                 260

ttt tgg cta gag gat gaa acg aga ttt ttt aaa atg cac tta gga cgg   870
Phe Trp Leu Glu Asp Glu Thr Arg Phe Phe Lys Met His Leu Gly Arg
            265                 270                 275

tat tca gga aat gct ggt gat gca ttc cgg ggt ctc aaa aaa gaa gat   918
Tyr Ser Gly Asn Ala Gly Asp Ala Phe Arg Gly Leu Lys Lys Glu Asp
        280                 285                 290

aat caa aat gca atg cct ttt agc aca tca gat gtt gat aat gat ggg   966
Asn Gln Asn Ala Met Pro Phe Ser Thr Ser Asp Val Asp Asn Asp Gly
    295                 300                 305

tgt cgc cct gca tgc ctg gtc aat ggt cag tct gtg aag agc tgc agt   1014
Cys Arg Pro Ala Cys Leu Val Asn Gly Gln Ser Val Lys Ser Cys Ser
310                 315                 320                 325
```

```
cac ctc cat aac aag acc ggc tgg tgg ttt aac gag tgt ggt cta gca   1062
His Leu His Asn Lys Thr Gly Trp Trp Phe Asn Glu Cys Gly Leu Ala
                330                 335                 340

aat cta aat ggc att cat cac ttc tct gga aaa ttg ctt gca act gga   1110
Asn Leu Asn Gly Ile His His Phe Ser Gly Lys Leu Leu Ala Thr Gly
            345                 350                 355

att caa tgg ggc acg tgg acc aaa aac aac tca cct gtc aag att aaa   1158
Ile Gln Trp Gly Thr Trp Thr Lys Asn Asn Ser Pro Val Lys Ile Lys
        360                 365                 370

tct gtt tca atg aaa att aga aga atg tac aat cca tat ttt aaa       1203
Ser Val Ser Met Lys Ile Arg Arg Met Tyr Asn Pro Tyr Phe Lys
    375                 380                 385

taatctcatt taacattgta atgcaagttc tacaatgata atatattaaa gatttttaaa 1263

agtttatctt ttcacttagt gtttcaaaca tattaggcaa aatttaactg tagatggcat 1323

ttagatgtta tgagtttaat tagaaaactt caattttgta gtattctata aaagaaaaca 1383

tggcttattg tatgttttta cttctgacta tattaacaat atacaatgaa atttgtttca 1443

agtgaactac aacttgtctt cctaaaattt atagtgattt taaaggattt tgccttttct 1503

ttgaagcatt tttaaaccat aatatgttgt aaggaaaatt gaagggaata ttttacttat 1563

ttttatactt tatatgatta tataatctac agataatttc tactgaagac agttacaata 1623

aataacttta tgcagattaa tatataagct acacatgatg taaaaacctt actatttcta 1683

ggtgatgcca taccatttta aaagtagtaa gagtttgctg cccaaatagt ttttcttgtt 1743

ttcatatcta atcatggtta actatttttgt tattgtttgt aataaatata tgtactttta 1803

tatcctgaaa aaaaaaaaaa a                                           1824
```

<210> 8
<211> 388
<212> PRT
<213> Homo sapiens

<400> 8

```
Met Met Ser Pro Ser Gln Ala Ser Leu Leu Phe Leu Asn Val Cys Ile
  1               5                  10                  15

Phe Ile Cys Gly Glu Ala Val Gln Gly Asn Cys Val His His Ser Thr
            20                  25                  30

Asp Ser Ser Val Val Asn Ile Val Glu Asp Gly Ser Asn Ala Lys Asp
        35                  40                  45

Glu Ser Lys Ser Asn Asp Thr Val Cys Lys Glu Asp Cys Glu Glu Ser
    50                  55                  60

Cys Asp Val Lys Thr Lys Ile Thr Arg Glu Glu Lys His Phe Met Cys
65                  70                  75                  80

Arg Asn Leu Gln Asn Ser Ile Val Ser Tyr Thr Arg Ser Thr Lys Lys
                85                  90                  95
```

```
Leu Leu Arg Asn Met Met Asp Glu Gln Gln Ala Ser Leu Asp Tyr Leu
            100                 105                 110

Ser Asn Gln Val Asn Glu Leu Met Asn Arg Val Leu Leu Leu Thr Thr
        115                 120                 125

Glu Val Phe Arg Lys Gln Leu Asp Pro Phe Pro His Arg Pro Val Gln
    130                 135                 140

Ser His Gly Leu Asp Cys Thr Asp Ile Lys Asp Thr Ile Gly Ser Val
145                 150                 155                 160

Thr Lys Thr Pro Ser Gly Leu Tyr Ile Ile His Pro Glu Gly Ser Ser
                165                 170                 175

Tyr Pro Phe Glu Val Met Cys Asp Met Asp Tyr Arg Gly Gly Gly Trp
            180                 185                 190

Thr Val Ile Gln Lys Arg Ile Asp Gly Ile Ile Asp Phe Gln Arg Leu
        195                 200                 205

Trp Cys Asp Tyr Leu Asp Gly Phe Gly Asp Leu Leu Gly Glu Phe Trp
    210                 215                 220

Leu Gly Leu Lys Lys Ile Phe Tyr Ile Val Asn Gln Lys Asn Thr Ser
225                 230                 235                 240

Phe Met Leu Tyr Val Ala Leu Glu Ser Glu Asp Asp Thr Leu Ala Tyr
                245                 250                 255

Ala Ser Tyr Asp Asn Phe Trp Leu Glu Asp Glu Thr Arg Phe Phe Lys
            260                 265                 270

Met His Leu Gly Arg Tyr Ser Gly Asn Ala Gly Asp Ala Phe Arg Gly
        275                 280                 285

Leu Lys Lys Glu Asp Asn Gln Asn Ala Met Pro Phe Ser Thr Ser Asp
    290                 295                 300

Val Asp Asn Asp Gly Cys Arg Pro Ala Cys Leu Val Asn Gly Gln Ser
305                 310                 315                 320

Val Lys Ser Cys Ser His Leu His Asn Lys Thr Gly Trp Trp Phe Asn
                325                 330                 335

Glu Cys Gly Leu Ala Asn Leu Asn Gly Ile His His Phe Ser Gly Lys
            340                 345                 350

Leu Leu Ala Thr Gly Ile Gln Trp Gly Thr Trp Thr Lys Asn Asn Ser
        355                 360                 365

Pro Val Lys Ile Lys Ser Val Ser Met Lys Ile Arg Arg Met Tyr Asn
    370                 375                 380

Pro Tyr Phe Lys
385
```

<210> 9
<211> 395
<212> DNA
<213> Homo sapiens

<220>
<223> CG250

<220>
<221> CDS
<222> (21)..(272)

<400> 9
gtccggattg atcaagagcc gcc tgt cca gtg tta tgt agg ggc aac ggg cag 53
Ala Cys Pro Val Leu Cys Arg Gly Asn Gly Gln
1 5 10

tac tcc aag ggc cgt tgc ctg tgt ttc agc ggc tgt aag ggc acc gag 101
Tyr Ser Lys Gly Arg Cys Leu Cys Phe Ser Gly Cys Lys Gly Thr Glu
15 20 25

tgt gat gtg ccg act acc cag tgt att gac cca cag tgc ggg gga cgt 149
Cys Asp Val Pro Thr Thr Gln Cys Ile Asp Pro Gln Cys Gly Gly Arg
30 35 40

ggg att tgt atc atg ggc tct tgt gct tgc aac tca gga tac aaa gga 197
Gly Ile Cys Ile Met Gly Ser Cys Ala Cys Asn Ser Gly Tyr Lys Gly
45 50 55

gaa agt tgt gaa gaa gca cca aga tac att cca gag aag gaa aga aaa 245
Glu Ser Cys Glu Glu Ala Pro Arg Tyr Ile Pro Glu Lys Glu Arg Lys
60 65 70 75

aaa aaa aag gcc tcc aac ttg cat gtt taggagaaat ggccggacat 292
Lys Lys Lys Ala Ser Asn Leu His Val
80

agcctaccaa agtggtggtt cttcgggatt ttttggacaa ggggggtaac taggaacttt 352

tctttttac taaaacactg aggttcgggg ggtttatttt gan 395

<210> 10
<211> 84
<212> PRT
<213> Homo sapiens

<400> 10
Ala Cys Pro Val Leu Cys Arg Gly Asn Gly Gln Tyr Ser Lys Gly Arg
1 5 10 15

Cys Leu Cys Phe Ser Gly Cys Lys Gly Thr Glu Cys Asp Val Pro Thr
20 25 30

Thr Gln Cys Ile Asp Pro Gln Cys Gly Gly Arg Gly Ile Cys Ile Met
35 40 45

Gly Ser Cys Ala Cys Asn Ser Gly Tyr Lys Gly Glu Ser Cys Glu Glu
50 55 60

Ala Pro Arg Tyr Ile Pro Glu Lys Glu Arg Lys Lys Lys Lys Ala Ser
65 70 75 80

Asn Leu His Val

<210> 11
<211> 1526
<212> DNA
<213> Homo sapiens

<220>
<223> CG006alt3 with 827 nt from CG006alt2

<220>
<221> CDS
<222> (64)..(1347)

<400> 11
gtctaggtct gcttccagaa gaaaacagtt ccacgttgct tgaaattgaa aatcaagata 60

aaa atg ttc aca att aag ctc ctt ctt ttt att gtt cct cta gtt att 108
Met Phe Thr Ile Lys Leu Leu Leu Phe Ile Val Pro Leu Val Ile
1 5 10 15

tcc tcc aga att gat caa gac aat tca tca ttt gat tct cta tct cca 156
Ser Ser Arg Ile Asp Gln Asp Asn Ser Ser Phe Asp Ser Leu Ser Pro
20 25 30

gag cca aaa tca aga ttt gct atg tta gac gat gta aaa att tta gcc 204
Glu Pro Lys Ser Arg Phe Ala Met Leu Asp Asp Val Lys Ile Leu Ala
35 40 45

aat ggc ctc ctt cag ttg gga cat ggt ctt aaa gac ttt gtc cat aag 252
Asn Gly Leu Leu Gln Leu Gly His Gly Leu Lys Asp Phe Val His Lys
50 55 60

acg aag ggc caa att aat gac ata ttt caa aaa ctc aac ata ttt gat 300
Thr Lys Gly Gln Ile Asn Asp Ile Phe Gln Lys Leu Asn Ile Phe Asp
65 70 75

cag tct ttt tat gat cta tcg ctg caa acc agt gaa atc aaa gaa gaa 348
Gln Ser Phe Tyr Asp Leu Ser Leu Gln Thr Ser Glu Ile Lys Glu Glu
80 85 90 95

gaa aag gaa ctg aga aga act aca tat aaa cta caa gtc aaa aat gaa 396
Glu Lys Glu Leu Arg Arg Thr Thr Tyr Lys Leu Gln Val Lys Asn Glu
100 105 110

gag gta aag aat atg tca ctt gaa ctc aac tca aaa ctt gaa agc ctc 444
Glu Val Lys Asn Met Ser Leu Glu Leu Asn Ser Lys Leu Glu Ser Leu
115 120 125

cta gaa gaa aaa att cta ctt caa caa aaa gtg aaa tat tta gaa gag 492
Leu Glu Glu Lys Ile Leu Leu Gln Gln Lys Val Lys Tyr Leu Glu Glu
130 135 140

caa cta act aac tta att caa aat caa cct gga act cca gaa cac cca 540
Gln Leu Thr Asn Leu Ile Gln Asn Gln Pro Gly Thr Pro Glu His Pro
145 150 155

gaa gta act tca ctt aaa act ttt gta gaa aaa caa gat aat agc atc 588
Glu Val Thr Ser Leu Lys Thr Phe Val Glu Lys Gln Asp Asn Ser Ile
160 165 170 175

aaa gac ctt ctc cag acc gtg gaa gac caa tat aaa caa tta aac caa 636
Lys Asp Leu Leu Gln Thr Val Glu Asp Gln Tyr Lys Gln Leu Asn Gln
180 185 190

```
cag cat agt caa ata aaa gaa ata gaa aat cag ctc aga agg act agt   684
Gln His Ser Gln Ile Lys Glu Ile Glu Asn Gln Leu Arg Arg Thr Ser
            195                 200                 205

att caa gaa ccc aca gaa att tct cta tct tcc aag cca aga gca cca   732
Ile Gln Glu Pro Thr Glu Ile Ser Leu Ser Ser Lys Pro Arg Ala Pro
        210                 215                 220

aga act act ccc ttt ctt cag ttg aat gaa ata aga aat gta aaa cat   780
Arg Thr Thr Pro Phe Leu Gln Leu Asn Glu Ile Arg Asn Val Lys His
    225                 230                 235

gat ggc att cct gct gaa tgt acc acc att tat aac aga ggt gaa cat   828
Asp Gly Ile Pro Ala Glu Cys Thr Thr Ile Tyr Asn Arg Gly Glu His
240                 245                 250                 255

aca agt ggc atg tat gcc atc aga ccc agc aac tct caa gtt ttt cat   876
Thr Ser Gly Met Tyr Ala Ile Arg Pro Ser Asn Ser Gln Val Phe His
                260                 265                 270

gtc tac tgg gat gtt ata tca gga gaa ttt tgg ttg ggc cta gag aag   924
Val Tyr Trp Asp Val Ile Ser Gly Glu Phe Trp Leu Gly Leu Glu Lys
            275                 280                 285

ata tac tcc ata gtg aag caa tct aat tat gtt tta cga att gag ttg   972
Ile Tyr Ser Ile Val Lys Gln Ser Asn Tyr Val Leu Arg Ile Glu Leu
        290                 295                 300

gaa gac tgg aaa gac aac aaa cat tat att gaa tat tct ttt tac ttg   1020
Glu Asp Trp Lys Asp Asn Lys His Tyr Ile Glu Tyr Ser Phe Tyr Leu
    305                 310                 315

gga aat cac gaa acc aac tat acg cta cat cta gtt gcg att act ggc   1068
Gly Asn His Glu Thr Asn Tyr Thr Leu His Leu Val Ala Ile Thr Gly
320                 325                 330                 335

aat gtc ccc aat gca atc ccg gaa aac aaa gat ttg gtg ttt tct act   1116
Asn Val Pro Asn Ala Ile Pro Glu Asn Lys Asp Leu Val Phe Ser Thr
                340                 345                 350

tgg gat cac aaa gca aaa gga cac ttc aac tgt cca gag ggt tat tca   1164
Trp Asp His Lys Ala Lys Gly His Phe Asn Cys Pro Glu Gly Tyr Ser
            355                 360                 365

gga ggc tgg tgg tgg cat gat gag tgt gga gaa aac aac cta aat ggt   1212
Gly Gly Trp Trp Trp His Asp Glu Cys Gly Glu Asn Asn Leu Asn Gly
        370                 375                 380

aaa tat aac aaa cca aga gca aaa tct aag cca gag agg aga aga gga   1260
Lys Tyr Asn Lys Pro Arg Ala Lys Ser Lys Pro Glu Arg Arg Arg Gly
    385                 390                 395

tta tct tgg aag tct caa aat gga agg tta tac tct ata aaa tca acc   1308
Leu Ser Trp Lys Ser Gln Asn Gly Arg Leu Tyr Ser Ile Lys Ser Thr
400                 405                 410                 415

aaa atg ttg atc cat cca aca gat tca gaa agc ttt gaa tgaactgagg    1357
Lys Met Leu Ile His Pro Thr Asp Ser Glu Ser Phe Glu
                420                 425

caaatttaaa aggcaataat ttaaacatta acctcattcc aagttaatgt ggtctaataa 1417
```

tctggtatta aatccttaag agaaagcttg agaaatagat ttttttttatc ttaaagtcac 1477

tgtctattta agattaaaca tacaatcaca taaccttaaa aaaaaaaaa 1526

<210> 12
<211> 428
<212> PRT
<213> Homo sapiens

<400> 12

```
Met Phe Thr Ile Lys Leu Leu Leu Phe Ile Val Pro Leu Val Ile Ser
  1               5                  10                  15

Ser Arg Ile Asp Gln Asp Asn Ser Ser Phe Asp Ser Leu Ser Pro Glu
             20                  25                  30

Pro Lys Ser Arg Phe Ala Met Leu Asp Asp Val Lys Ile Leu Ala Asn
         35                  40                  45

Gly Leu Leu Gln Leu Gly His Gly Leu Lys Asp Phe Val His Lys Thr
     50                  55                  60

Lys Gly Gln Ile Asn Asp Ile Phe Gln Lys Leu Asn Ile Phe Asp Gln
 65                  70                  75                  80

Ser Phe Tyr Asp Leu Ser Leu Gln Thr Ser Glu Ile Lys Glu Glu Glu
                 85                  90                  95

Lys Glu Leu Arg Arg Thr Thr Tyr Lys Leu Gln Val Lys Asn Glu Glu
            100                 105                 110

Val Lys Asn Met Ser Leu Glu Leu Asn Ser Lys Leu Glu Ser Leu Leu
        115                 120                 125

Glu Glu Lys Ile Leu Leu Gln Gln Lys Val Lys Tyr Leu Glu Glu Gln
    130                 135                 140

Leu Thr Asn Leu Ile Gln Asn Gln Pro Gly Thr Pro Glu His Pro Glu
145                 150                 155                 160

Val Thr Ser Leu Lys Thr Phe Val Glu Lys Gln Asp Asn Ser Ile Lys
                165                 170                 175

Asp Leu Leu Gln Thr Val Glu Asp Gln Tyr Lys Gln Leu Asn Gln Gln
            180                 185                 190

His Ser Gln Ile Lys Glu Ile Glu Asn Gln Leu Arg Arg Thr Ser Ile
        195                 200                 205

Gln Glu Pro Thr Glu Ile Ser Leu Ser Ser Lys Pro Arg Ala Pro Arg
    210                 215                 220

Thr Thr Pro Phe Leu Gln Leu Asn Glu Ile Arg Asn Val Lys His Asp
225                 230                 235                 240

Gly Ile Pro Ala Glu Cys Thr Thr Ile Tyr Asn Arg Gly Glu His Thr
                245                 250                 255

Ser Gly Met Tyr Ala Ile Arg Pro Ser Asn Ser Gln Val Phe His Val
            260                 265                 270
```

```
Tyr Trp Asp Val Ile Ser Gly Glu Phe Trp Leu Gly Leu Glu Lys Ile
        275                 280                 285

Tyr Ser Ile Val Lys Gln Ser Asn Tyr Val Leu Arg Ile Glu Leu Glu
    290                 295                 300

Asp Trp Lys Asp Asn Lys His Tyr Ile Glu Tyr Ser Phe Tyr Leu Gly
305                 310                 315                 320

Asn His Glu Thr Asn Tyr Thr Leu His Leu Val Ala Ile Thr Gly Asn
                325                 330                 335

Val Pro Asn Ala Ile Pro Glu Asn Lys Asp Leu Val Phe Ser Thr Trp
            340                 345                 350

Asp His Lys Ala Lys Gly His Phe Asn Cys Pro Glu Gly Tyr Ser Gly
        355                 360                 365

Gly Trp Trp Trp His Asp Glu Cys Gly Glu Asn Asn Leu Asn Gly Lys
    370                 375                 380

Tyr Asn Lys Pro Arg Ala Lys Ser Lys Pro Glu Arg Arg Arg Gly Leu
385                 390                 395                 400

Ser Trp Lys Ser Gln Asn Gly Arg Leu Tyr Ser Ile Lys Ser Thr Lys
                405                 410                 415

Met Leu Ile His Pro Thr Asp Ser Glu Ser Phe Glu
            420                 425
```

<210> 13
<211> 1836
<212> DNA
<213> Homo sapiens

<220>
<223> CG0015 genomic sequence

<220>
<223> Exons are located from nucleotides 1-196, 382-567,
573-652, 831-1104, 1674-1836

<400> 13

```
tgatgttcag gaggaaagcg aggtctccag cttttctagg aactcatctt gtacctcatt 60

atccgctcag cgcgcagccg cccggttcca ccagctggac gtcaagttcc gcgagctggc 120

gcagctcgtc acccagcaga gcagtctcat cgcccgcctg gagcgcctgt gcccgggagg 180

cgcgggcggg cagcagnnnn natgttggtc aggctggtct cgaactcctg acctcaggtg 240

atccgcccgc ctcggcctcc caaagtgctg ggattacagg cacgggccat agcgcccagc 300

ctgtctgcac tttaaagcca agttgtttag cttttgggga ggatcattcc tagggctggg 360

acacccccac tgccagatgt ccaggtcctg ccgctacccc cactggtgcc tgtggttccg 420

gtccgtcttg tgggtagcac cagtgacacc agtaggatgc tggacccagc cccagagccc 480

cagagagacc agacccagag acagcaggag cccatggctt ctcccatgcc tgcaggtcac 540
```

cctgcggtcc ccaccaagcc tgtgggcnnn nnccgtggca ggattgtgca gaggcccgcc 600

aggcaggcca tgaacagagt ggagtgtatg aactgcgagt gggccgtcac gtagtgtcag 660

tatggtgtga gcagcaactg gagggtggag gctggactgt gatccagcgg aggcaagatg 720

gttcagtcaa cttcttcact acctggcagc actataaggt gggcacaggt gggcagaggc 780

agggaagggg agggagcctg ttctggcttc ctgactttcc tgccctgcca ggcgggcttt 840

gggcggccag acggagaata ctggctgggc cttgaacccg tgtatcagct gaccagccgt 900

ggggaccatg agctgctggt tctcctggag gactgggggg gccgtggagc acgtgcccac 960

tatgatggct tctccctgga acccgagagc gaccactacc gcctgcggct tggccagtac 1020

catggtgatg ctggagactc tctttcctgg cacaatgaca agcccttcag caccgtggat 1080

agggaccgag actcctattc tggtaaggag aactcctatt ctggtgagag gataggggag 1140

gcgggactcc tgttctggtg agggaatgaa aggaggtagg gtaggtaaga cgcccctctg 1200

gtaagtataa ggataagcaa gcttttattc cgtcaagaga acaaaggtca ggacttttat 1260

cctggtgggg ggatggggag tccasattcc ttctstgatg aggcaaaaaa agaatcaaga 1320

ctcctgttca agtawagggc agagggtgag agctagtact cttattctag aaaggaagta 1380

gatacttttc tttgataaag gaatgaacgg tagactccta gtttgcagaa aaggtgggaa 1440

agatgtgact tgtactttgg taaggagata gggaaggaat taaggctatt actctgaaga 1500

aagttggggg gccagggctc ctatttttt gctgaggaga tggaagatca gggcttgtat 1560

tcaataagaa tgggaggggc caggggatgc ctggcaaaag ccttgcactg tgaggtgcag 1620

gtagaggctt ttattctggt gagaggacat ggactctctc tctcccctca ggtaactgtg 1680

ccctgtacca gcggggaggc tggtggtacc atgcctgtgc ccactccaac ctcaacggtg 1740

tgtggcacca cggcggccac taccgaagcc gctaccagga tggtgtctac tgggctgagt 1800

ttcgtggtgg ggcatattct ctcaggaagg ccgcca 1836

<210> 14
<211> 2319
<212> DNA
<213> Homo sapiens

<220>
<223> CG06 genomic

<220>
<223> Exons are located from nucleotides 1-368,
671-1201, 1207-1302, 1308-1572, 1970-2319

<400> 14
gtctaggtct gcttccagaa gaaaacagtt ccacgttgct tgaaattgaa aatcaagata 60

aaaatgttca caattaagct ccttcttttt attgttcctc tagttatttc ctccagaatt 120

gatcaagaca attcatcatt tgattctcta tctccagagc caaaatcaag atttgctatg 180

ttagacgatg taaaaatttt agccaatggc ctccttcagt tgggacatgg tcttaaagac 240

tttgtccata agacgaaggg ccaaattaat gacatatttc aaaaactcaa catatttgat 300

cagtcttttt atgatctatc gctgcaaacc agtgaaatca aagaagaaga aaaggaactg 360

agaagaacnn nnntctgcat cctatggaac aaggcacaga atttaatgtt caattgcaag 420

ataaaactca ggaaaaatat gaaagggtat aacttttaaa tcaaatttca gttatgagaa 480

tcgatcaata ctaccaccct gtgtttgact cagcttttga agaggccaaa gagaaatgac 540

taaatgtcac ttcctgttac accacgttcc agtgacccag tggtgaccgt aggaccagca 600

caaacatgtg tttcacatct gcagtgagtc tccgctctaa gccaccactc cttatctggc 660

tatagggcct tacatataaa ctacaagtca aaaatgaaga ggtaaagaat atgtcacttg 720

aactcaactc aaaacttgaa agcctcctag aagaaaaaat tctacttcaa caaaaagtga 780

aatatttaga agagcaacta actaacttaa ttcaaaatca acctggaact ccagaacacc 840

cagaagtaac ttcacttaaa acttttgtag aaaaacaaga taatagcatc aaagaccttc 900

tccagaccgt ggaagaccaa tataaacaat taaaccaaca gcatagtcaa ataaaagaaa 960

tagaaaatca gctcagaagg actagtattc aagaacccac agaaatttct ctatcttcca 1020

agccaagagc accaagaact actccctttc ttcagttgaa tgaaataaga aatgtaaaac 1080

atgatggcat tcctgctgaa tgtaccacca tttataacag aggtgaacat acaagtggca 1140

tgtatgccat cagacccagc aactctcaag tttttcatgt ctactgtgat gttatatcag 1200

gnnnnntagt ccatggacat taattcaaca tcgaatagat ggatcacaaa acttcaatga 1260

aacgtgggag aactacaaat atggttttgg gaggcttgat ggnnnnnaga attttggttg 1320

ggcctagaga agatatactc tatagtgaag caatctaatt atgttttacg aattgagttg 1380

gaagactgga aagacaacaa acattatatt gaatattctt tttacttggg aaatcacgaa 1440

accaactata cgctacatct agttgcgatt actggcaatg tccccaatgc aatcccggaa 1500

aacaaagatt tggtgttttc tacttgggat cacaaagcaa aaggacactt caactgtcca 1560

gagggttatt caggtatctt tttctgatac caatacttta ttttcatatc ttcaaagtat 1620

cttcccacat tattagctat tatctgcaat gacaactttt aaaaatccga atcccaaata 1680

agcgttttct ctctagacga aaacctctta actataatga aagtgttcat tctagttcaa 1740

tcaggtattt tacctctaat cttcctcaga ttttctattt tttggtagtg tatagattat 1800

ttatacagat tatttaaaat tgggacttat acagattatt taaaactggg atacatgcat 1860

ctaaaacact gtaatattta taagaaagga agataaactt acggggaaat acagtaacag 1920

taactacata cgagtctgta cccattaaat tgcatatcta tctcctttag gaggctggtg 1980

gtggcatgat gagtgtggag aaaacaacct aaatggtaaa tataacaaac caagagcaaa 2040

atctaagcca gagaggagaa gaggattatc ttggaagtct caaaatggaa ggttatactc 2100

tataaaatca accaaaatgt tgatccatcc aacagattca gaaagctttg aatgaactga 2160

ggcaaattta aaaggcaata atttaaacat taacctcatt ccaagttaat gtggtctaat 2220

aatctggtat taaatcctta agagaaagct tgagaaatag attttttta tcttaaagtc 2280

actgtctatt taagattaaa catacaatca cataacctt 2319

<210> 15
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer pSPORT VP1

<400> 15
aggcacccca ggctttacac ttta 24

<210> 16
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer pSPORT VP2

<400> 16
ttcccgggtc gacgatttcg t 21

<210> 17
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer Marathon VP1

<400> 17
ccatcctaat acgactcact atagggc 27

<210> 18
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer Marathon VP2

<400> 18
actcactata gggctcgagc ggc 23

<210> 19
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG006R5

<400> 19
gtctttccag tcttccaact caattcgta 29

<210> 20
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG0006R6

<400> 20
gtatatcttc tctaggccca a 21

<210> 21
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG0006R11

<400> 21
gatgttgaat taatgtccat ggactacctg at 32

<210> 22
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG0006R10

<400> 22
ggcatacatg ccacttgtat gtt 23

<210> 23
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG0006R12

<400> 23
gattttgaat taagttagtt agttgctctt ctaaa 35

<210> 24
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG0006R13

<400> 24
gagttgagtt caagtgacat a 21

<210> 25
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG0006R15

<400> 25
tcattaattt ggcccttcgt cttatggaca aa 32

<210> 26
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG0006R16

<400> 26
gtcccaactg aaggaggcca t 21

<210> 27
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG0007R1

<400> 27
gcaggctata tgccgtgttc tcgccacca 29

<210> 28
<211> 21

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG0007R2

<400> 28
cccgcagttg cacggccagg c 21

<210> 29
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG007R5

<400> 29
tgctgaattc gcaggtgctg ctt 23

<210> 30
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG007R6

<400> 30
gctgggccac cttgtgga 18

<210> 31
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG007R7

<400> 31
ctgcaggagt ccgtgcgcca ggacatt 27

<210> 32
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG007R8

<400> 32
atctcgtccc aggacgcaaa 20

<210> 33
<211> 30
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG144R1

<400> 33
ccatgtgact gaacaggtct gtgaggaaaa 30

<210> 34
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG144R2

<400> 34
gaactctatt catgagctcg tta 23

<210> 35
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG144R3

<400> 35
acatgattcc tcacagtctt ccttacaaa 29

<210> 36
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG144R4

<400> 36
actactgaag agtccgtaga a 21

<210> 37
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG015R1

<400> 37
gaaagagagt ctccagcatc acctaccat 29

<210> 38
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG015R3

<400> 38
ccagggagaa gccatcatag t 21

<210> 39
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG015alt1R1

<400> 39
ggctctgggg ctgggtccag catccta 27

<210> 40
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG015alt1R6

<400> 40
acccacaaga cggaccggaa 20

<210> 41
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG015alt2R5

<400> 41
gggtgacctg caggcatggg agaagcat 28

<210> 42
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG015alt2R6

<400> 42
ggctgggtcc agcatccta 19

<210> 43
<211> 27
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG015alt1R5

<400> 43
ggctctgggg ctgggtccag catccta 27

<210> 44
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG015alt1R7

<400> 44
gtggcggcag gacctgct 18

<210> 45
<211> 1139
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (3)..(839)

<220>
<223> CG015alt1

<400> 45

```
tg atg ttc agg agg aaa gcg agg tct cca gct ttt cta gga act cat    47
   Met Phe Arg Arg Lys Ala Arg Ser Pro Ala Phe Leu Gly Thr His
     1               5                  10                  15

ctt gta cct cat tat ccg ctc agc gcg cag ccg ccc ggt tcc acc agc   95
Leu Val Pro His Tyr Pro Leu Ser Ala Gln Pro Pro Gly Ser Thr Ser
                 20                  25                  30

tgg acg tca agt tcc gcg agc tgg cgc agc tcg tca ccc agc aga gca   143
Trp Thr Ser Ser Ser Ala Ser Trp Arg Ser Ser Ser Pro Ser Arg Ala
             35                  40                  45

gtc tca tcg ccc gcc tgg agc gcc tgt gcc cgg gag gcg cgg gcg ggc   191
Val Ser Ser Pro Ala Trp Ser Ala Cys Ala Arg Glu Ala Arg Ala Gly
         50                  55                  60

agc agc agg tcc tgc cgc cac ccc cac tgg tgc ctg tgg ttc cgg tcc   239
Ser Ser Arg Ser Cys Arg His Pro His Trp Cys Leu Trp Phe Arg Ser
     65                  70                  75

gtc ttg tgg gta gca cca gtg aca cca gta gga tgc tgg acc cag ccc   287
Val Leu Trp Val Ala Pro Val Thr Pro Val Gly Cys Trp Thr Gln Pro
 80                  85                  90                  95
```

```
cag agc ccc aga gag acc aga ccc aga gac agc agg agc cca tgg ctt   335
Gln Ser Pro Arg Glu Thr Arg Pro Arg Asp Ser Arg Ser Pro Trp Leu
                100                 105                 110

ctc cca tgc ctg cag gtc acc ctg cgg tcc cca cca agc ctg tgg gcg   383
Leu Pro Cys Leu Gln Val Thr Leu Arg Ser Pro Pro Ser Leu Trp Ala
            115                 120                 125

ggc ttt ggg cgg cca gac gga gaa tac tgg ctg ggc ctt gaa ccc gtg   431
Gly Phe Gly Arg Pro Asp Gly Glu Tyr Trp Leu Gly Leu Glu Pro Val
        130                 135                 140

tat cag ctg acc agc cgt ggg gac cat gag ctg ctg gtt ctc ctg gag   479
Tyr Gln Leu Thr Ser Arg Gly Asp His Glu Leu Leu Val Leu Leu Glu
    145                 150                 155

gac tgg ggg ggc cgt gga gca cgt gcc cac tat gat ggc ttc tcc ctg   527
Asp Trp Gly Gly Arg Gly Ala Arg Ala His Tyr Asp Gly Phe Ser Leu
160                 165                 170                 175

gaa ccc gag agc gac cac tac cgc ctg cgg ctt ggc cag tac cat ggt   575
Glu Pro Glu Ser Asp His Tyr Arg Leu Arg Leu Gly Gln Tyr His Gly
                180                 185                 190

gat gct gga gac tct ctt tcc tgg cac aat gac aag ccc ttc agc acc   623
Asp Ala Gly Asp Ser Leu Ser Trp His Asn Asp Lys Pro Phe Ser Thr
            195                 200                 205

gtg gat agg gac cga gac tcc tat tct ggt aac tgt gcc ctg tac cag   671
Val Asp Arg Asp Arg Asp Ser Tyr Ser Gly Asn Cys Ala Leu Tyr Gln
        210                 215                 220

cgg gga ggc tgg tgg tac cat gcc tgt gcc cac tcc aac ctc aac ggt   719
Arg Gly Gly Trp Trp Tyr His Ala Cys Ala His Ser Asn Leu Asn Gly
    225                 230                 235

gtg tgg cac cac ggc ggc cac tac cga agc cgc tac cag gat ggt gtc   767
Val Trp His His Gly Gly His Tyr Arg Ser Arg Tyr Gln Asp Gly Val
240                 245                 250                 255

tac tgg gct gag ttt cgt ggt ggg gca tat tct ctc agg aag gcc gcc   815
Tyr Trp Ala Glu Phe Arg Gly Gly Ala Tyr Ser Leu Arg Lys Ala Ala
                260                 265                 270

atg ctc att cgg ccc ctg aag ctg tgactctgtg ttcctctgtc ccctaggccc  869
Met Leu Ile Arg Pro Leu Lys Leu
            275

tagaggacat tggtcagcag gagcccaagt tgttctggcc acaccttctt tgtggctcag 929

tgccaatgtg tcccacagaa cttcccactg tggatctgtg accctgggcg ctgaaaatgg 989

gacccaggaa tccccccccgt caatatcttg gcctcagatg gctccccaag gtcattcata 1049

tctcggtttg agctcatatc ttataataac acaaagtagc cacagaccgt gtctggtttg 1109

tatctgcacc tggcaggggt cactccctgg                                  1139
```

<210> 46
<211> 279
<212> PRT

<213> Homo sapiens

<400> 46

```
Met Phe Arg Arg Lys Ala Arg Ser Pro Ala Phe Leu Gly Thr His Leu
  1               5                  10                  15

Val Pro His Tyr Pro Leu Ser Ala Gln Pro Pro Gly Ser Thr Ser Trp
             20                  25                  30

Thr Ser Ser Ser Ala Ser Trp Arg Ser Ser Ser Pro Ser Arg Ala Val
         35                  40                  45

Ser Ser Pro Ala Trp Ser Ala Cys Ala Arg Glu Ala Arg Ala Gly Ser
     50                  55                  60

Ser Arg Ser Cys Arg His Pro His Trp Cys Leu Trp Phe Arg Ser Val
 65                  70                  75                  80

Leu Trp Val Ala Pro Val Thr Pro Val Gly Cys Trp Thr Gln Pro Gln
                 85                  90                  95

Ser Pro Arg Glu Thr Arg Pro Arg Asp Ser Arg Ser Pro Trp Leu Leu
            100                 105                 110

Pro Cys Leu Gln Val Thr Leu Arg Ser Pro Pro Ser Leu Trp Ala Gly
        115                 120                 125

Phe Gly Arg Pro Asp Gly Glu Tyr Trp Leu Gly Leu Glu Pro Val Tyr
    130                 135                 140

Gln Leu Thr Ser Arg Gly Asp His Glu Leu Leu Val Leu Leu Glu Asp
145                 150                 155                 160

Trp Gly Gly Arg Gly Ala Arg Ala His Tyr Asp Gly Phe Ser Leu Glu
                165                 170                 175

Pro Glu Ser Asp His Tyr Arg Leu Arg Leu Gly Gln Tyr His Gly Asp
            180                 185                 190

Ala Gly Asp Ser Leu Ser Trp His Asn Asp Lys Pro Phe Ser Thr Val
        195                 200                 205

Asp Arg Asp Arg Asp Ser Tyr Ser Gly Asn Cys Ala Leu Tyr Gln Arg
    210                 215                 220

Gly Gly Trp Trp Tyr His Ala Cys Ala His Ser Asn Leu Asn Gly Val
225                 230                 235                 240

Trp His His Gly Gly His Tyr Arg Ser Arg Tyr Gln Asp Gly Val Tyr
                245                 250                 255

Trp Ala Glu Phe Arg Gly Gly Ala Tyr Ser Leu Arg Lys Ala Ala Met
            260                 265                 270

Leu Ile Arg Pro Leu Lys Leu
        275
```

<210> 47
<211> 1280
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (2)..(1027)

<220>
<223> CG015alt2

<400> 47
```
t gat gtt cag gag gaa agc gag gtc tcc agc ttt tct agg aac tca tct 49
  Asp Val Gln Glu Glu Ser Glu Val Ser Ser Phe Ser Arg Asn Ser Ser
    1               5                  10                  15

tgt acc tca tta tcc gct cag cgc gca gcc gcc cgg ttc cac cag ctg   97
Cys Thr Ser Leu Ser Ala Gln Arg Ala Ala Ala Arg Phe His Gln Leu
            20                  25                  30

gac gtc aag ttc cgc gag ctg gcg cag ctc gtc acc cag cag agc agt   145
Asp Val Lys Phe Arg Glu Leu Ala Gln Leu Val Thr Gln Gln Ser Ser
        35                  40                  45

ctc atc gcc cgc ctg gag cgc ctg tgc ccg gga ggc gcg ggc ggg cag   193
Leu Ile Ala Arg Leu Glu Arg Leu Cys Pro Gly Gly Ala Gly Gly Gln
    50                  55                  60

cag cag gtc ctg ccg cta ccc cca ctg gtg cct gtg gtt ccg gtc cgt   241
Gln Gln Val Leu Pro Leu Pro Pro Leu Val Pro Val Val Pro Val Arg
65                  70                  75                  80

ctt gtg ggt agc acc agt gac acc agt agg atg ctg gac cca gcc cca   289
Leu Val Gly Ser Thr Ser Asp Thr Ser Arg Met Leu Asp Pro Ala Pro
                85                  90                  95

gag ccc cag aga gac cag acc cag aga cag cag gag ccc atg gct tct   337
Glu Pro Gln Arg Asp Gln Thr Gln Arg Gln Gln Glu Pro Met Ala Ser
            100                 105                 110

ccc atg cct gca ggt cac cct gcg gtc ccc acc aag cct gtg ggc ccg   385
Pro Met Pro Ala Gly His Pro Ala Val Pro Thr Lys Pro Val Gly Pro
        115                 120                 125

tgg cag gat tgt gca gag gcc cgc cag gca ggc cat gaa cag agt gga   433
Trp Gln Asp Cys Ala Glu Ala Arg Gln Ala Gly His Glu Gln Ser Gly
    130                 135                 140

gtg tat gaa ctg cga gtg ggc cgt cac gta gtg tca gta tgg tgt gag   481
Val Tyr Glu Leu Arg Val Gly Arg His Val Val Ser Val Trp Cys Glu
145                 150                 155                 160

cag caa ctg gag ggt gga ggc tgg act gtg atc cag cgg agg caa gat   529
Gln Gln Leu Glu Gly Gly Gly Trp Thr Val Ile Gln Arg Arg Gln Asp
                165                 170                 175

ggt tca gtc aac ttc ttc act acc tgg cag cac tat aag gcg ggc ttt   577
Gly Ser Val Asn Phe Phe Thr Thr Trp Gln His Tyr Lys Ala Gly Phe
            180                 185                 190

ggg cgg cca gac gga gaa tac tgg ctg ggc ctt gaa ccc gtg tat cag   625
Gly Arg Pro Asp Gly Glu Tyr Trp Leu Gly Leu Glu Pro Val Tyr Gln
        195                 200                 205
```

```
ctg acc agc cgt ggg gac cat gag ctg ctg gtt ctc ctg gag gac tgg   673
Leu Thr Ser Arg Gly Asp His Glu Leu Leu Val Leu Leu Glu Asp Trp
    210                 215                 220

ggg ggc cgt gga gca cgt gcc cac tat gat ggc ttc tcc ctg gaa ccc   721
Gly Gly Arg Gly Ala Arg Ala His Tyr Asp Gly Phe Ser Leu Glu Pro
225                 230                 235                 240

gag agc gac cac tac cgc ctg cgg ctt ggc cag tac cat ggt gat gct   769
Glu Ser Asp His Tyr Arg Leu Arg Leu Gly Gln Tyr His Gly Asp Ala
                245                 250                 255

gga gac tct ctt tcc tgg cac aat gac aag ccc ttc agc acc gtg gat   817
Gly Asp Ser Leu Ser Trp His Asn Asp Lys Pro Phe Ser Thr Val Asp
            260                 265                 270

agg gac cga gac tcc tat tct ggt aac tgt gcc ctg tac cag cgg gga   865
Arg Asp Arg Asp Ser Tyr Ser Gly Asn Cys Ala Leu Tyr Gln Arg Gly
        275                 280                 285

ggc tgg tgg tac cat gcc tgt gcc cac tcc aac ctc aac ggt gtg tgg   913
Gly Trp Trp Tyr His Ala Cys Ala His Ser Asn Leu Asn Gly Val Trp
    290                 295                 300

cac cac ggc ggc cac tac cga agc cgc tac cag gat ggt gtc tac tgg   961
His His Gly Gly His Tyr Arg Ser Arg Tyr Gln Asp Gly Val Tyr Trp
305                 310                 315                 320

gct gag ttt cgt ggt ggg gca tat tct ctc agg aag gcc gcc atg ctc   1009
Ala Glu Phe Arg Gly Gly Ala Tyr Ser Leu Arg Lys Ala Ala Met Leu
                325                 330                 335

att cgg ccc ctg aag ctg tgactctgtg ttcctctgtc ccctaggccc          1057
Ile Arg Pro Leu Lys Leu
            340

tagaggacat tggtcagcag gagcccaagt tgttctggcc acaccttctt tgtggctcag 1117

tgccaatgtg tcccacagaa cttcccactg tggatctgtg accctgggcg ctgaaaatgg 1177

gacccaggaa tccccccccgt caatatcttg gcctcagatg gctccccaag gtcattcata 1237

tctcggtttg agctcatatc ttataataac acaaagtagc cac                   1280


<210> 48
<211> 342
<212> PRT
<213> Homo sapiens

<400> 48
Asp Val Gln Glu Glu Ser Glu Val Ser Ser Phe Ser Arg Asn Ser Ser
  1               5                  10                  15

Cys Thr Ser Leu Ser Ala Gln Arg Ala Ala Ala Arg Phe His Gln Leu
             20                  25                  30

Asp Val Lys Phe Arg Glu Leu Ala Gln Leu Val Thr Gln Gln Ser Ser
         35                  40                  45

Leu Ile Ala Arg Leu Glu Arg Leu Cys Pro Gly Gly Ala Gly Gly Gln
     50                  55                  60
```

```
Gln Gln Val Leu Pro Leu Pro Pro Leu Val Pro Val Val Pro Val Arg
 65                  70                  75                  80

Leu Val Gly Ser Thr Ser Asp Thr Ser Arg Met Leu Asp Pro Ala Pro
                 85                  90                  95

Glu Pro Gln Arg Asp Gln Thr Gln Arg Gln Gln Glu Pro Met Ala Ser
            100                 105                 110

Pro Met Pro Ala Gly His Pro Ala Val Pro Thr Lys Pro Val Gly Pro
        115                 120                 125

Trp Gln Asp Cys Ala Glu Ala Arg Gln Ala Gly His Glu Gln Ser Gly
    130                 135                 140

Val Tyr Glu Leu Arg Val Gly Arg His Val Val Ser Val Trp Cys Glu
145                 150                 155                 160

Gln Gln Leu Glu Gly Gly Gly Trp Thr Val Ile Gln Arg Arg Gln Asp
                165                 170                 175

Gly Ser Val Asn Phe Phe Thr Thr Trp Gln His Tyr Lys Ala Gly Phe
            180                 185                 190

Gly Arg Pro Asp Gly Glu Tyr Trp Leu Gly Leu Glu Pro Val Tyr Gln
        195                 200                 205

Leu Thr Ser Arg Gly Asp His Glu Leu Leu Val Leu Leu Glu Asp Trp
    210                 215                 220

Gly Gly Arg Gly Ala Arg Ala His Tyr Asp Gly Phe Ser Leu Glu Pro
225                 230                 235                 240

Glu Ser Asp His Tyr Arg Leu Arg Leu Gly Gln Tyr His Gly Asp Ala
                245                 250                 255

Gly Asp Ser Leu Ser Trp His Asn Asp Lys Pro Phe Ser Thr Val Asp
            260                 265                 270

Arg Asp Arg Asp Ser Tyr Ser Gly Asn Cys Ala Leu Tyr Gln Arg Gly
        275                 280                 285

Gly Trp Trp Tyr His Ala Cys Ala His Ser Asn Leu Asn Gly Val Trp
    290                 295                 300

His His Gly Gly His Tyr Arg Ser Arg Tyr Gln Asp Gly Val Tyr Trp
305                 310                 315                 320

Ala Glu Phe Arg Gly Gly Ala Tyr Ser Leu Arg Lys Ala Ala Met Leu
                325                 330                 335

Ile Arg Pro Leu Lys Leu
            340
```

<210> 49
<211>
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer CG007 (Forward) #72

<400> 49
gaaagatctg ggacccgtgc agtccaagtc

<210> 50
<211>
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG007 (Reverse) #73

<400> 50
gcagcagagg cagcctccta agaagagctc gagaagcttg ac

<210> 51
<211>
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG144 (Forward) #90

<400> 51
gattgatcaa gtacaaggta actgtgtaca tc

<210> 52
<211>
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
CG144 (Reverse) #91

<400> 52
gtacaatcca tattttaaat aatctcgagc catggatc